(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 August 2001 (23.08.2001)

PCT

(10) International Publication Number
WO 01/60860 A2

(51) International Patent Classification[7]: C07K 14/47, 16/18, C12Q 1/68, C07K 14/705, G01N 33/50

(21) International Application Number: PCT/US01/05171

(22) International Filing Date: 20 February 2001 (20.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/183,319	17 February 2000 (17.02.2000)	US
60/189,862	16 March 2000 (16.03.2000)	US
60/207,454	25 May 2000 (25.05.2000)	US
60/211,314	9 June 2000 (09.06.2000)	US
60/219,007	18 July 2000 (18.07.2000)	US
60/255,281	13 December 2000 (13.12.2000)	US

(71) Applicant: MILLENNIUM PREDICTIVE MEDICINE, INC. [US/US]; One Kendall Square Building 700, Cambridge, MA 02139 (US).

(72) Inventors: SCHLEGEL, Robert; 211 Melrose Street, Auburndale, MA 02466 (US). ENDEGE, Wilson, O.; 222 Normandy Drive, Norwood, MA 02062 (US). MONAHAN, John, E.; 942 West Street, Walpole, MA 02081 (US).

(74) Agents: SMITH, DeAnn, F. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/60860 A2

(54) Title: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF HUMAN PROSTATE CANCER

(57) Abstract: The invention relates to compositions, kits, and methods for detecting, characterizing, preventing, and treating human prostate cancers. A variety of novel markers are provided, wherein changes in the levels of expression of one or more of the markers is correlated with the presence of prostate cancer.

NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF HUMAN PROSTATE CANCER

RELATED APPLICATIONS

The present application claims priority to U.S. provisional patent application serial no. 60/183,319, filed on February 17, 2000, U.S. provisional patent application serial no. 60/189,862, filed on March 16, 2000, U.S. provisional patent application serial no. 60/207,454, filed on May 25, 2000, U.S. provisional patent application serial no. 60/211,314, filed on June 9, 2000 U.S. provisional patent application serial no. 60/219,007, filed on July 18, 2000, and U.S. provisional patent application serial no. 60/255,281, filed on December 13, 2000, all of which are expressly incorporated by reference.

FIELD OF THE INVENTION

The field of the invention is prostate cancer, including diagnosis, characterization, management, and therapy of prostate cancer.

BACKGROUND OF THE INVENTION

The increased number of cancer cases reported in the United States, and, indeed, around the world, is a major concern. Currently there are only a handful of treatments available for specific types of cancer, and these provide no absolute guarantee of success. In order to be most effective, these treatments require not only an early detection of the malignancy, but also a reliable assessment of the severity of the malignancy.

Carcinoma of the prostate (PCA) is the most frequently diagnosed cancer in men in the United States, and is the second leading cause of male cancer deaths (Karp *et al.*, 1996, *Cancer Res.* 56:5547-5556). The acute susceptibility of this organ to cancer in men is not understood. Skenes glands represent a tissue in females that is homologous to the male prostate, but not a site where significant neoplastic transformation is observed.

An unusual challenge presented by prostate cancer is that most prostate tumors do not represent life threatening conditions. Projections from autopsy surveys indicate

that as many as 11 million American men have prostate cancer (Dhom, 1983, *J. Cancer Res. Clin. Oncol.*, 106:210-218). These figures are consistent with clinical observations of prostate carcinomas, which normally exhibit a slow and lingering course of progression. Such disease progression results in relatively few prostate tumors developing into cases of clinical concern during the lifetime of the patient. If, upon detection with available methods, the cancer appears well-differentiated, organ-confined and focal, treatment normally can not extend the life expectancy of older patients.

Unfortunately, the prostate carcinomas that are progressive in nature frequently have already metastasized by the time of clinical detection with available methods. Survival rates for individuals with metastatic prostate cancer are quite low. Between these two extremes are patients with prostate tumors that will metastasize during their lifetimes, but have not yet done so. For these patients, surgical removal of the prostate is curative and extends life expectancy. Therefore, accurate determination of which group a newly diagnosed patient falls into is critical in determining optimal treatment and patient survival.

Currently there is at least one early and noninvasive test available to the physician for detecting asymptomatic disease. The presence of Prostate Specific Antigen (PSA) can be measured with relative ease from blood samples using standard antibody-based detection kits. Abnormally high levels of this antigen in a patient's serum indicate a likelihood of prostate disease, possibly either a carcinoma, Benign Prostatic Hyperplasia (BPH) or prostatitis. In the majority of cases, PSA elevation is due to BPH or prostatitis rather than carcinoma.

Although clinical and pathologic stage and histological grading systems (e.g., Gleason's) have been used to indicate prognosis for groups of patients based on the degree of tumor differentiation or the type of glandular pattern (Carter and Coffey, In: J. P. Karr and H. Yamanak (eds.), *Prostate Cancer: The Second Tokyo Symposium*, pp. 19-27, New York: Elsevier, 1989.; Diamond *et al.*, *J. Urol.*, 128: 729-734, 1982), these systems do not adequately predict the progression rate of the cancer. While the use of computer-system image analysis of histologic sections of primary lesions for "nuclear roundness" has been suggested as an aide in the management of individual patients (Diamond *et al.*, 1982, *J. Urol.*, 128:729-734), this method is of limited use in studying the progression of the disease.

The analysis of DNA content/ploidy using flow cytometry and FISH has been demonstrated to have utility predicting prostate cancer aggressiveness (Pearsons *et al.*, 1993, *J. Urol.*, 150:120-125; Macoska *et al.*, 1994, *Cancer Res.*, 54: 3824-3830; Visakorpi *et al.*, 1994, *Am. J. Pathol.*, 145:1-7; Takahashi *et al.*, 1994, *Cancer Res.*, 54:3574-3579; Alcaraz *et al.*, *Cancer Res.*, 55:3998-4002, 1994), but these methods are expensive, time-consuming, and the latter methodology requires the construction of centromere-specific probes for analysis. There also exist specific nuclear matrix proteins whose expression has been reported to be associated with prostate cancer. However, these protein markers apparently do not distinguish between BPH and prostate cancer (Partin *et al.*, 1993, *Cancer Res.*, 53:744-746). Unfortunately, markers that cannot distinguish between benign and malignant prostate tumors are of little value.

It would therefore be beneficial to provide specific methods and reagents for the diagnosis, staging, prognosis, monitoring, and treatment of diseases associated with prostate cancer, or to indicate a predisposition to such for preventative medicine.

SUMMARY OF THE INVENTION

The invention relates to novel genes associated with prostate cancer as well as methods of assessing whether a patient is afflicted with or has higher than normal risk for developing prostate cancer. The method of the present invention comprises the step of comparing the level of expression of a marker (listed within Tables 1-9) in a patient sample with the normal level of expression of the marker in a control, *e.g.*, a sample from a patient without prostate cancer. A significant difference between the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with prostate cancer or has higher than normal risk for developing prostate cancer.

In one embodiment of the methods of the present invention, the sample comprises cells obtained from the patient. The cells may be found in a prostate tissue sample collected, for example, by a prostate tissue biopsy or histology section, or a bone marrow biopsy. In another embodiment, the patient sample is a prostate-associated body fluid. Such fluids include, for example, blood fluids, lymph, urine, prostatic fluid and semen.

In accordance with the methods of the present invention, the presence and/or level of expression of the marker in a sample can be assessed, for example, by detecting the presence in the sample of:

- a protein or protein fragment corresponding to the marker (*e.g.* using a reagent, such as an antibody, an antibody derivative, or an antibody fragment, which binds specifically with the protein or protein fragment)
- a transcribed polynucleotide (*e.g.* an mRNA or a cDNA), or fragment thereof, having at least a portion with which the marker is substantially homologous (*e.g.* by contacting a mixture of transcribed polynucleotides obtained from the sample with a substrate having one or more of the markers listed within Tables 1-9 fixed thereto at selected positions)
- a metabolite which is produced directly (*i.e.*, catalyzed) or indirectly by a protein corresponding to the marker
- a transcribed polynucleotide or fragment thereof, wherein the polynucleotide anneals with the marker under stringent hybridization conditions.

The methods of the present invention are useful for further diagnosing patients having an identified prostate mass or symptoms associated with prostate cancer, *e.g.* abnormally high levels of PSA. The methods of the present invention can further be of particular use with patients having an enhanced risk of developing prostate cancer (*e.g.*, patients having a familial history of prostate cancer and patients identified as having a mutant oncogene). The methods of the present invention may further be of particular use in monitoring the efficacy of treatment of a prostate cancer patient (*e.g.* the efficacy of chemotherapy).

All cancers have staging schemes that are used to describe the degree to which the cancer has progressed. The TNM staging approach assigns the primary tumor (T) to one of four stages (and to additional substages within these categories) based on the size and location of the primary tumor within the prostate. A T1 designation indicates a microscopic tumor which cannot be detected by a digital rectal exam. A T2NO designation refers to a tumor palpable upon a digital rectal exam but are contained within the prostate capsule (local disease). In all forms of stage T3 disease the tumors have extended through the prostate capsule into the surrounding connective tissue or seminal vesicles. The T4 designation refers to tumors that have escaped from the

prostate and can be found in the pelvic region. The N stage refers to whether the primary tumor has spread to the regional lymph nodes (pelvic lymph nodes). The M stage refers to whether the tumor cells have metastasized to distant sites.

The methods of the present invention may be performed using a plurality (*e.g.* 2, 3, 5, or 10 or more) of markers. According to a method involving a plurality of markers, the level of expression in the sample of each of a plurality of markers independently selected from the markers listed in Tables 1-9 is compared with the normal level of expression of each of the plurality of markers in samples of the same type obtained from control humans not afflicted with prostate cancer. A significantly altered level of expression in the sample of one or more of the markers listed in Tables 1-9, or some combination thereof, relative to that marker's corresponding normal levels, is an indication that the patient is afflicted with prostate cancer. The markers of Tables 1-9 may also be used in combination with known prostate cancer markers in the methods of the present invention, *e.g.* PSA analysis.

In a preferred method of assessing whether a patient is afflicted with prostate cancer (*e.g.*, new detection ("screening"), detection of recurrence, reflex testing), the method comprises comparing:

a) the level of expression of a marker in a patient sample, wherein at least one marker is selected from the markers of Tables 1-9, and

b) the normal level of expression of the marker in a control non-prostate cancer sample.

A significant difference between the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with prostate cancer.

The invention further relates to a method of assessing the efficacy of a therapy for inhibiting prostate cancer in a patient. This method comprises comparing:

a) expression of a marker in a first sample obtained from the patient prior to providing at least a portion of the therapy to the patient, wherein the marker is selected from the group consisting of the markers listed within Tables 1-9, and

b) expression of the marker in a second sample obtained from the patient following provision of the portion of the therapy.

A significant difference between the level of expression of the marker in the second sample, relative to the first sample, is an indication that the therapy is efficacious for inhibiting prostate cancer in the patient.

It will be appreciated that in this method the "therapy" may be any therapy for treating prostate cancer including, but not limited to, chemotherapy, immunotherapy, gene therapy, radiation therapy and surgical removal of tissue. Thus, the methods of the invention may be used to evaluate a patient before, during and after therapy, for example, to evaluate the reduction in tumor burden.

The present invention therefore further comprises a method for monitoring the progression of prostate cancer in a patient, the method comprising:

a) detecting in a patient sample at a first time point, the expression of a marker, wherein the marker is selected from the group consisting of the markers listed in Tables 1-9;

b) repeating step a) at a subsequent time point in time; and

c) comparing the level of expression detected in steps a) and b), and therefrom monitoring the progression of prostate cancer in the patient.

The present invention also includes a method for assessing the aggressiveness or indolence of prostate cancer (*e.g.*, staging), the method comprising comparing:

a) the level of expression of a marker in a patient sample, wherein at least one marker is selected from the markers of Tables 1-9, and

b) the normal level of expression of the marker in a control sample.

A significant difference between the level of expression in the sample and the normal level is an indication that the cancer is aggressive or indolent.

The present invention further includes a method for determining whether prostate cancer has metastasized or is likely to metastasize in the future, the method comprising comparing:

a) the level of expression of a marker in a patent sample, wherein at least one marker is selected from the markers of Tables 1-9 and

b) the normal level (or non-metastatic level) of expression of the marker in a control sample.

A significant difference between the level of expression in the patient sample and the normal level (or non-metastatic level) is an indication that the prostate cancer has metastasized or is likely to metastasize in the future.

The invention also includes a method of selecting a composition for inhibiting prostate cancer in a patient. This method comprises the steps of:

a) obtaining a sample comprising cancer cells from the patient;

b) separately maintaining aliquots of the sample in the presence of a plurality of test compositions;

c) comparing expression of a marker listed within Tables 1-9 in each of the aliquots; and

d) selecting one of the test compositions which alters the level of expression of the marker in the aliquot containing that test composition, relative to other test compositions.

In addition, the invention includes a method of inhibiting prostate cancer in a patient. This method comprises the steps of:

a) obtaining a sample comprising cancer cells from the patient;

b) separately maintaining aliquots of the sample in the presence of a plurality of test compositions;

c) comparing expression of a marker listed within Tables 1-9 in each of the aliquots; and

d) administering to the patient at least one of the test compositions which alters the level of expression of the marker in the aliquot containing that test composition, relative to other test compositions.

The invention also includes a kit for assessing whether a patient is afflicted with prostate cancer. This kit comprises reagents for assessing expression of a marker listed within Tables 1-9.

In another aspect, the invention relates to a kit for assessing the suitability of each of a plurality of compounds for inhibiting a prostate cancer in a patient. The kit comprises a reagent for assessing expression of a marker listed within Tables 1-9, and may also comprise a plurality of compounds.

In another aspect, the invention relates to a kit for assessing the presence of prostate cancer cells. This kit comprises an antibody, wherein the antibody binds

specifically with a protein or protein fragment corresponding to a marker listed within Tables 1-9. The kit may also comprise a plurality of antibodies, wherein the plurality binds specifically with a protein or protein fragment corresponding to a different marker listed within Tables 1-9.

The invention also includes a kit for assessing the presence of prostate cancer cells, wherein the kit comprises a nucleic acid probe. The probe binds specifically with a transcribed polynucleotide corresponding to a marker listed within Tables 1-9. The kit may also comprise a plurality of probes, wherein each of the probes binds specifically with a transcribed polynucleotide corresponding to a different marker listed within Tables 1-9.

The invention further relates to a method of making an isolated hybridoma which produces an antibody useful for assessing whether a patient is afflicted with prostate cancer. The method comprises isolating a protein or protein fragment corresponding to a marker listed within Tables 1-9, immunizing a mammal using the isolated protein or protein fragment, isolating splenocytes from the immunized mammal, fusing the isolated splenocytes with an immortalized cell line to form hybridomas, and screening individual hybridomas for production of an antibody which specifically binds with the protein or protein fragment, to isolate the hybridoma. The invention also includes an antibody produced by this method.

The invention further includes a method of assessing the prostate carcinogenic potential of a test compound. This method comprises the steps of:

a) maintaining separate aliquots of prostate cells in the presence and absence of the test compound; and

b) comparing expression of a marker in each of the aliquots.

The marker is selected from those listed within Tables 1-9. A significant difference between the level of expression of the marker in the aliquot maintained in the presence of (or exposed to) the test compound, relative to the aliquot maintained in the absence of the test compound, is an indication that the test compound possesses prostate carcinogenic potential.

Additionally, the invention includes a kit for assessing the prostate carcinogenic potential of a test compound. The kit comprises prostate cells and a reagent for

assessing expression of a marker in each of the aliquots. The marker is selected from those listed within Tables 1-9.

The invention further relates to a method of treating a patient afflicted with prostate cancer. This method comprises providing to cells of the patient an antisense oligonucleotide complementary to a polynucleotide corresponding to a marker listed within Tables 1-9, which is overexpressed in prostate cancer. In an alternative method, expression of a gene corresponding to a marker selected from the markers listed in Tables 1-9 which is underexpressed in prostate cancer, is increased.

The invention includes a method of inhibiting prostate cancer in a patient at risk for developing prostate cancer. This method comprises inhibiting or increasing expression (or overexpression) of a gene corresponding to a marker listed within Tables 1-9, that is either overexpressed or underexpressed, respectively, in prostate cancer.

It will be appreciated that the methods and kits of the present invention may also include known cancer markers including known prostate cancer markers. It will further be appreciated that the methods and kits may be used to identify cancers other than prostate cancer.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to newly discovered genes associated with the cancerous state of prostate cells. It has been discovered that the level of expression of these individual genes, also referred to as markers, and combinations of these genes, correlates with the presence of prostate cancer or a pre-malignant condition in a patient. Methods are provided for detecting the presence of prostate cancer in a sample, the absence of prostate cancer in a sample, the stage of a prostate cancer, the metastatic potential of a prostate cancer, the indolence or aggressiveness of the cancer, and other characteristics of prostate cancer that are relevant to prevention, diagnosis, characterization and therapy of prostate cancer in a patient.

Definitions

As used herein, each of the following terms has the meaning associated with it in this section.

The articles "a" and "an" are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

A "marker" is a naturally-occurring polymer corresponding to at least one of the novel nucleic acids listed within Tables 1-9. For example, markers include, without limitation, sense and anti-sense strands of genomic DNA (*i.e.* including any introns occurring therein), RNA generated by transcription of genomic DNA (*i.e.* prior to splicing), RNA generated by splicing of RNA transcribed from genomic DNA, and proteins generated by translation of spliced RNA (*i.e.* including proteins both before and after cleavage of normally cleaved regions such as transmembrane signal sequences). As used herein, "marker" may also include a cDNA made by reverse transcription of an RNA generated by transcription of genomic DNA (including spliced RNA).

As used herein a polynucleotide "corresponds to" another (a first) polynucleotide if it is related to the first polynucleotide by any of the following relationships: The second polynucleotide comprises the first polynucleotide and the second polynucleotide encodes a gene product; 2) The second polynucleotide is 5' or 3' to the first polynucleotide in cDNA, RNA, genomic DNA, or fragment of any of these polynucleotides. For example, a second polynucleotide may be a fragment of a gene that includes the first and second polynucleotides. The first and second polynucleotides are related in that they are components of the gene coding for a gene product, such as a protein or antibody. However, it is not necessary that the second polynucleotide comprises or overlaps with the first polynucleotide to be encompassed within the definition of "corresponding to" as used herein. For example, the first polynucleotide may be a fragment of a 3' untranslated region of the second polynucleotide. The first and second polynucleotide may be fragments of a gene coding for a gene product. The second polynucleotide may be an exon of the gene while the first polynucleotide may be an intron of the gene; 3) The second polynucleotide is the complement of the first polynucleotide.

The term "probe" refers to any molecule which is capable of selectively binding to a specifically intended target molecule, for example a marker of the invention. Probes can be either synthesized by one skilled in the art, or derived from appropriate biological preparations. For purposes of detection of the target molecule, probes may be

specifically designed to be labeled, as described herein. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, cDNA, proteins, antibodies, and organic monomers.

A "prostate-associated" body fluid is a fluid which, when in the body of a patient, contacts or passes through prostate cells or into which cells or proteins shed from prostate cells are capable of passing. Exemplary prostate-associated body fluids include blood fluids, semen, prostate fluid, lymph and urine.

The "normal" level of expression of a marker is the level of expression of the marker in prostate cells or prostate-associated body fluids of a patient, *e.g.* a human, not afflicted with prostate cancer.

"Over-expression" and "under-expression" of a marker refer to expression of the marker of a patient at a greater or lesser level, respectively, than normal level of expression of the marker (*e.g.* at least two-fold greater or lesser level).

As used herein, the term "promoter/regulatory sequence" means a nucleic acid sequence which is required for expression of a gene product operably linked to the promoter/regulatory sequence. In some instances, this sequence may be the core promoter sequence and in other instances, this sequence may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product. The promoter/regulatory sequence may, for example, be one which expresses the gene product in a tissue-specific manner.

A "constitutive" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell under most or all physiological conditions of the cell.

An "inducible" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell substantially only when an inducer which corresponds to the promoter is present in the cell.

A "tissue-specific" promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living human cell substantially only if the cell is a cell of the tissue type corresponding to the promoter.

A "transcribed polynucleotide" is a polynucleotide (*e.g.* an RNA, a cDNA, or an analog of one of an RNA or cDNA) which is complementary to or homologous with all or a portion of a mature RNA made by transcription of a genomic DNA corresponding to a marker of the invention and normal post-transcriptional processing (*e.g.* splicing), if any, of the transcript.

"Complementary" refers to the broad concept of sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of forming specific hydrogen bonds ("base pairing") with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. More preferably, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

"Homologous" as used herein, refers to nucleotide sequence similarity between two regions of the same nucleic acid strand or between regions of two different nucleic acid strands. Homology between two regions is expressed in terms of the proportion of nucleotide residue positions of the two regions that are occupied by the same nucleotide residue. By way of example, a region having the nucleotide sequence 5'-ATTGCC-3' and a region having the nucleotide sequence 5'-TATGGC-3' share 50% homology. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residue positions of each of the

portions are occupied by the same nucleotide residue. More preferably, all nucleotide residue positions of each of the portions are occupied by the same nucleotide residue.

A marker is "fixed" to a substrate if it is covalently or non-covalently associated with the substrate such that the substrate can be rinsed with a fluid (*e.g.* standard saline citrate, pH 7.4) without a substantial fraction of the marker dissociating from the substrate.

As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature.

Expression of a marker in a patient is "significantly" higher than the normal level of expression of a marker if the level of expression of the marker is greater than the normal level by an amount greater than the standard error of the assay employed to assess expression, and preferably at least twice, and more preferably three, four, five or ten times that amount. Alternately, expression of the marker in the patient can be considered "significantly" higher or lower than the normal level of expression if the level of expression is at least about two, and preferably at least about three, four, or five times, higher or lower, respectively, than the normal level of expression of the marker.

Prostate cancer is "inhibited" if at least one symptom of the cancer is alleviated, terminated, slowed, or prevented. As used herein, prostate cancer is also "inhibited" if recurrence or metastasis of the cancer is reduced, slowed, delayed, or prevented.

A kit is any manufacture (*e.g.* a package or container) comprising at least one reagent, *e.g.* a probe, for specifically detecting a marker of the invention, the manufacture being promoted, distributed, or sold as a unit for performing the methods of the present invention.

Description

The present invention is based, in part, on identification of novel markers which are differentially expressed in prostate cancer cells when compared with normal (*i.e.* non- cancerous) prostate cells. The markers of the invention correspond to DNA, RNA, and polypeptide molecules which can be detected in one or both of normal and cancerous prostate cells. The presence, absence, or level of expression of one or more of these markers in prostate cells is herein correlated with the cancerous state of the tissue. The invention thus includes compositions, kits, and methods for assessing the

cancerous state of prostate cells (*e.g.* cells obtained from a human, cultured human cells, archived or preserved human cells and *in vivo* cells).

The compositions, kits, and methods of the invention have the following uses, among others:

1) assessing whether a patient is afflicted with prostate cancer;
2) assessing the stage of prostate cancer in a human patient;
3) assessing the grade of prostate cancer in a patient;
4) assessing the benign or malignant nature of prostate cancer in a patient;
5) assessing the metastatic potential of prostate cancer in a patient;
6) assessing the histological type of neoplasm (*e.g. Adenocarcinoma*) associated with prostate cancer in a patient;
7) assessing the indolent or aggressive nature of prostate cancer in a patient;
8) making an isolated hybridoma which produces an antibody useful for assessing whether a patient is afflicted with prostate cancer;
9) assessing the presence of prostate cancer cells;
10) assessing the efficacy of one or more test compounds for inhibiting prostate cancer in a patient;
11) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
12) monitoring the progression of prostate cancer in a patient;
13) selecting a composition or therapy for inhibiting prostate cancer in a patient;
14) treating a patient afflicted with prostate cancer;
15) inhibiting prostate cancer in a patient;
16) assessing the prostate carcinogenic potential of a test compound; and
17) inhibiting prostate cancer in a patient at risk for developing prostate cancer.

The invention thus includes a method of assessing whether a patient is afflicted with prostate cancer which includes assessing whether the patient has pre-metastasized prostate cancer. This method comprises comparing the level of expression of a marker in a patient sample and the normal level of expression of the marker in a control, *e.g.*, a

non-prostate cancer sample. A significant difference between the level of expression of the marker in the patient sample and the normal level is an indication that the patient is afflicted with prostate cancer. The marker is selected from the group consisting of the markers listed within Tables 1-9. Although one or more molecules corresponding to the markers listed within Tables 1-9 may have been described by others, the significance of the level of expression of these markers with regard to the cancerous state of prostate cells has not previously been recognized.

The polynucleotides set forth in Tables 1-9 represent previously unidentified nucleotide sequences. These nucleotide sequences were identified through subtracted library experiments described herein. Also provided by this invention are polynucleotides that correspond to the polynucleotides of Tables 1-9. In one embodiment, these polynucleotides are obtained by identification of a larger fragment or full-length coding sequence of these polynucleotides. Gene delivery vehicles, host cells, compositions and databases (all described herein) containing these polynucleotides are also provided by this invention.

The invention also encompasses polynucleotides which differ from that of the polynucleotides described herein, but which produce the same phenotypic effect, such as an allelic variant. These altered, but phenotypically equivalent polynucleotides are referred to as "equivalent nucleic acids." This invention also encompasses polynucleotides characterized by changes in non-coding regions that do not alter the polypeptide produced therefrom when compared to the polynucleotide herein. This invention further encompasses polynucleotides, which hybridize to the polynucleotides of the subject invention under conditions of moderate or high stringency. Alternatively, the polynucleotides are at least 85%, or at least 90%, or more preferably, greater or equal to 95% identical as determined by a sequence alignment program when run under default parameters.

Any marker or combination of markers listed within Tables 1-9, as well as any known markers in combination with the markers set forth within Tables 1-9, may be used in the compositions, kits, and methods of the present invention. In general, it is preferable to use markers for which the difference between the level of expression of the marker in prostate cancer cells or prostate-associated body fluids and the level of expression of the same marker in normal prostate cells or prostate-associated body fluids

is as great as possible. Although this difference can be as small as the limit of detection of the method for assessing expression of the marker, it is preferred that the difference be at least greater than the standard error of the assessment method, and preferably a difference of at least 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 15-, 20-, 25-, 100-, 500-, 1000-fold or greater.

It will be appreciated that patient samples containing prostate cells may be used in the methods of the present invention. In these embodiments, the level of expression of the marker can be assessed by assessing the amount (*e.g.* absolute amount or concentration) of the marker in a prostate cell sample, *e.g.*, prostate tissue sample obtained from a patient. The cell sample can, of course, be subjected to a variety of well-known post-collection preparative and storage techniques (*e.g.* fixation, storage, freezing, lysis, homogenization, DNA or RNA extraction, ultrafiltration, concentration, evaporation, centrifugation, etc.) prior to assessing the amount of the marker in the sample.

It will also be appreciated that certain markers correspond to proteins which are secreted from prostate cells (*i.e.* one or both of normal and cancerous cells) to the extracellular space surrounding the cells. These markers are preferably used in certain embodiments of the compositions, kits, and methods of the invention, owing to the fact that the protein corresponding to each of these markers can be detected in a prostate-associated body fluid sample. In addition, preferred *in vivo* techniques for detection of a protein corresponding to a marker of the invention include introducing into a subject a labeled antibody directed against the protein. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

Although not every marker corresponding to a secreted protein is indicated as such herein, it is a simple matter for the skilled artisan to determine whether any particular marker corresponds to a secreted protein. In order to make this determination, the protein corresponding to a marker is expressed in a test cell (*e.g.* a cell of a prostate cell line), extracellular fluid is collected, and the presence or absence of the protein in the extracellular fluid is assessed (*e.g.* using a labeled antibody which binds specifically with the protein).

The following is an example of a method which can be used to detect secretion of a protein corresponding to a marker of the invention. About 8 x 10^5 293T cells are incubated at 37°C in wells containing growth medium (Dulbecco's modified Eagle's medium {DMEM} supplemented with 10% fetal bovine serum) under a 5% (v/v) CO_2, 95% air atmosphere to about 60-70% confluence. The cells are then transfected using a standard transfection mixture comprising 2 micrograms of DNA comprising an expression vector encoding the protein and 10 microliters of LipofectAMINE™ (GIBCO/BRL Catalog no. 18342-012) per well. The transfection mixture is maintained for about 5 hours, and then replaced with fresh growth medium and maintained in an air atmosphere. Each well is gently rinsed twice with DMEM which does not contain methionine or cysteine (DMEM-MC; ICN Catalog no. 16-424- 54). About 1 milliliter of DMEM-MC and about 50 microcuries of Trans-^{35}S™ reagent (ICN Catalog no. 51006) are added to each well. The wells are maintained under the 5% CO_2 atmosphere described above and incubated at 37°C for a selected period. Following incubation, 150 microliters of conditioned medium is removed and centrifuged to remove floating cells and debris. The presence of the protein in the supernatant is an indication that the protein is secreted.

Examples of prostate-associated body fluids include blood fluids (*e.g.* whole blood, blood serum, blood having platelets removed therefrom, lymph, urine, prostatic fluid and semen. Many prostate-associated body fluids (*i.e.* usually excluding urine) can have prostate cells therein, particularly when the prostate cells are cancerous, and, more particularly, when the prostate cancer is metastasizing. Cell-containing fluids which can contain prostate cancer cells include, but are not limited to, whole blood, blood having platelets removed therefrom, lymph, prostatic fluid, and semen. Thus, the compositions, kits, and methods of the invention can be used to detect expression of markers corresponding to proteins having at least one portion which is displayed on the surface of cells which express it. Although the proteins having at least one cell-surface portion are not set forth herein, it is a simple matter for the skilled artisan to determine whether the protein corresponding to any particular marker comprises a cell-surface protein. For example, immunological methods may be used to detect such proteins on whole cells, or well known computer-based sequence analysis methods (*e.g.* the SIGNALP program;

Nielsen *et al.*, 1997, *Protein Engineering* 10:1-6) may be used to predict the presence of at least one extracellular domain (*i.e.* including both secreted proteins and proteins having at least one cell-surface domain). Expression of a marker corresponding to a protein having at least one portion which is displayed on the surface of a cell which expresses it may be detected without necessarily lysing the cell (*e.g.* using a labeled antibody which binds specifically with a cell-surface domain of the protein).

Expression of a marker of the invention may be assessed by any of a wide variety of well known methods for detecting expression of a transcribed molecule or protein. Non-limiting examples of such methods include immunological methods for detection of secreted, cell-surface, cytoplasmic, or nuclear proteins, protein purification methods, protein function or activity assays, nucleic acid hybridization methods, nucleic acid reverse transcription methods, and nucleic acid amplification methods.

In another preferred embodiment, expression of a marker is assessed using an antibody (*e.g.* a radio-labeled, chromophore-labeled, fluorophore-labeled, or enzyme-labeled antibody), an antibody derivative (*e.g.* an antibody conjugated with a substrate or with the protein or ligand of a protein-ligand pair {*e.g.* biotin-streptavidin}), or an antibody fragment (*e.g.* a single-chain antibody, an isolated antibody hypervariable domain, etc.) which binds specifically with a protein or protein fragment corresponding to the marker, such as the protein encoded by the open reading frame corresponding to the marker or such a protein which has undergone all or a portion of its normal post-translational modification.

In another preferred embodiment, expression of a marker is assessed by preparing mRNA/cDNA (*i.e.* a transcribed polynucleotide) from cells in a patient sample, and by hybridizing the mRNA/cDNA with a reference polynucleotide which is a complement of a polynucleotide comprising the marker, and fragments thereof. cDNA can, optionally, be amplified using any of a variety of polymerase chain reaction methods prior to hybridization with the reference polynucleotide. Expression of one or more markers can likewise be detected using quantitative PCR to assess the level of expression of the marker(s). Alternatively, any of the many known methods of detecting mutations or variants (*e.g.* single nucleotide polymorphisms, deletions, etc.) of a marker of the invention may be used to detect occurrence of a marker in a patient.

In a related embodiment, a mixture of transcribed polynucleotides obtained from the sample is contacted with a substrate having fixed thereto a polynucleotide complementary to or homologous with at least a portion (*e.g.* at least 7, 10, 15, 20, 25, 30, 40, 50, 100, 500, or more nucleotide residues) of a marker of the invention. If polynucleotides complementary to or homologous with a marker of the invention are differentially detectable on the substrate (*e.g.* detectable using radioactivity, different chromophores or fluorophores), are fixed to different selected positions, then the levels of expression of a plurality of markers can be assessed simultaneously using a single substrate (*e.g.* a "gene chip" microarray of polynucleotides fixed at selected positions). When a method of assessing marker expression is used which involves hybridization of one nucleic acid with another, it is preferred that the hybridization be performed under stringent hybridization conditions.

Because the compositions, kits, and methods of the invention rely on detection of a difference in expression levels of one or more markers of the invention, it is preferable that the level of expression of the marker is significantly greater than the minimum detection limit of the method used to assess expression in at least one of normal prostate cells and cancerous prostate cells.

It is understood that by routine screening of additional patient samples using one or more of the markers of the invention, it will be realized that certain of the markers are over- or underexpressed in cancers of various types, including specific prostate cancers, as well as other cancers such as breast or ovarian cancers. For example, it will be confirmed that some of the markers of the invention are over-expressed in most (*i.e.* 50% or more) or substantially all (*i.e.* 80% or more) of prostate cancer. Furthermore, it will be confirmed that certain of the markers of the invention are associated with prostate cancer of various stages.

It will be appreciated that as a greater number of patient samples are assessed for expression of the markers of the invention and the outcomes of the individual patients from whom the samples were obtained are correlated, it will also be confirmed that altered expression of certain of the markers of the invention are strongly correlated with malignant cancers and that altered expression of other markers of the invention are strongly correlated with benign tumors. The compositions, kits, and methods of the invention are thus useful for characterizing one or more of the stage, grade, histological

type, metastatic potential, indolent vs. aggressive phenotype and benign/malignant nature of prostate cancer in patients.

When the compositions, kits, and methods of the invention are used for characterizing one or more of the stage, grade, histological type, metastatic potential, indolent vs. aggressive phenotype and benign/malignant nature of prostate cancer in a patient, it is preferred that the marker or panel of markers of the invention is selected such that a positive result is obtained in at least about 20%, and preferably at least about 40%, 60%, or 80%, and more preferably in substantially all patients afflicted with a prostate cancer of the corresponding stage, grade, histological type, metastatic potential, indolent vs. aggressive phenotype or benign/malignant nature. Preferably, the marker or panel of markers of the invention is selected such that a positive predictive value (PPV) of greater than about 10% is obtained for the general population.

When a plurality of markers of the invention are used in the compositions, kits, and methods of the invention, the level of expression of each marker in a patient sample can be compared with the normal level of expression of each of the plurality of markers in non-cancerous samples of the same type, either in a single reaction mixture (*i.e.* using reagents, such as different fluorescent probes, for each marker or a mixture of similiarly labeled probes to access a plurality of markers that are fixed to a single substrate at different positions) or in individual reaction mixtures corresponding to one or more of the markers. In one embodiment, a significantly enhanced level of expression of more than one of the plurality of markers in the sample, relative to the corresponding normal levels, is an indication that the patient is afflicted with prostate cancer. When a plurality of markers is used, it is preferred that 2, 3, 4, 5, 8, 10, 12, 15, 20, 30, or 50 or more individual markers be used, wherein fewer markers are preferred.

In order to maximize the sensitivity of the compositions, kits, and methods of the invention (*i.e.* by interference attributable to cells of non-prostate origin in a patient sample), it is preferable that the marker of the invention used therein be a marker which has a restricted tissue distribution, *e.g.*, normally not expressed in non-prostate tissue.

Only a small number of markers are known to be associated with prostate cancers (*e.g.* PSA, PSMA, PAP , PCA3, PCTA-1, PSCA and STEAP). These markers are not, of course, included among the markers of the invention, although they may be used together with one or more markers of the invention in a panel of markers, for

example. It is well known that certain types of genes, such as oncogenes, tumor suppressor genes, growth factor-like genes, protease-like genes, and protein kinase-like genes are often involved with development of cancers of various types. Thus, among the markers of the invention, use of those which correspond to proteins which resemble known proteins encoded by known oncogenes and tumor suppressor genes, and those which correspond to proteins which resemble growth factors, proteases, and protein kinases are preferred.

Known oncogenes and tumor suppressor genes include, for example, *abl, abr, akT2NO, apc, bcl2α, bcl2β, bcl3, bcr, brca1, brca2, cbl, ccnd1, cdc42, cdk4, crk- II, csf1r/fms, dbl, dcc, dpc4/smad4, e-cad, e2f1/rbap, egfr/erbb-1, elk1, elk3, eph, erg, ets1, ets2, fer, fgr/src2, fli1/ergb2, fos, fps/fes, fra1, fra2, fyn, hck, hek, her2/erbb- 2/neu, her3/erbb-3, her4/erbb-4, hras1, hsT2NO, hstf1, igfbp2, ink4a, ink4b, inT2NO/fgf3, jun, junb, jund, kip2, kit, kras2a, kras2b, lck, lyn, mas, max, mcc, mdm2, met, mlh1, mmp10, mos, msh2, msh3, msh6, myb, myba, mybb, myc, mycl1, mycn, nf1, nf2, nme2, nras, p53, pdgfb, phb, pim1, pms1, pms2, ptc, pten, raf1, rap1a, rb1, rel, ret, ros1, ski, src1, tal1, tgfbr2, tgfb3, tgfbr3, thra1, thrb, tiam1, timp3, tjp1, tp53, trk, vav, vhl, vil2, waf1, wnt1, wnT2NO, wt1,* and *yes1* (Hesketh, 1997, In: *The Oncogene and Tumour Suppressor Gene Facts Book*, 2nd Ed., Academic Press; Fishel *et al.*, 1994, *Science* 266:1403-1405).

Known growth factors include platelet-derived growth factor alpha, platelet-derived growth factor beta (simian sarcoma viral {v-sis} oncogene homolog), thrombopoietin (myeloproliferative leukemia virus oncogene ligand, megakaryocyte growth and development factor), erythropoietin, B cell growth factor, macrophage stimulating factor 1 (hepatocyte growth factor-like protein), hepatocyte growth factor (hepapoietin A), insulin-like growth factor 1 (somatomedia C), hepatoma-derived growth factor, amphiregulin (schwannoma-derived growth factor), bone morphogenetic proteins 1, 2, 3, 3 beta, and 4, bone morphogenetic protein 7 (osteogenic protein 1), bone morphogenetic protein 8 (osteogenic protein 2), connective tissue growth factor, connective tissue activation peptide 3, epidermal growth factor (EGF), teratocarcinoma-derived growth factor 1, endothelin, endothelin 2, endothelin 3, stromal cell-derived factor 1, vascular endothelial growth factor (VEGF), VEGF-B, VEGF-C, placental growth factor (vascular endothelial growth factor-related protein), transforming growth

factor alpha, transforming growth factor beta 1 and its precursors, transforming growth factor beta 2 and its precursors, fibroblast growth factor 1 (acidic), fibroblast growth factor 2 (basic), fibroblast growth factor 5 and its precursors, fibroblast growth factor 6 and its precursors, fibroblast growth factor 7 (keratinocyte growth factor), fibroblast growth factor 8 (androgen-induced), fibroblast growth factor 9 (glia-activating factor), pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1), brain-derived neurotrophic factor, and recombinant glial growth factor 2.

Known proteases include interleukin-1 beta convertase and its precursors, Mch6 and its precursors, Mch2 isoform alpha, Mch4, Cpp32 isoform alpha, Lice2 gamma cysteine protease, Ich-1S, Ich-1L, Ich-2 and its precursors, TY protease, matrix metalloproteinase 1 (interstitial collagenase), matrix metalloproteinase 2 (gelatinase A, 72kD gelatinase, 72kD type IV collagenase), matrix metalloproteinase 7 (matrilysin), matrix metalloproteinase 8 (neutrophil collagenase), matrix metalloproteinase 12 (macrophage elastase), matrix metalloproteinase 13 (collagenase 3), metallopeptidase 1, cysteine-rich metalloprotease (disintegrin) and its precursors, subtilisin-like protease Pc8 and its precursors, chymotrypsin, snake venom-like protease, cathepsin I, cathepsin D (lysosomal aspartyl protease), stromelysin, aminopeptidase N, plasminogen, tissue plasminogen activator, plasminogen activator inhibitor type II, and urokinase-type plasminogen activator.

Known protein kinases include DAP kinase, serine/threonine protein kinases NIK, PK428, Krs-2, SAK, and EMK, interferon-inducible double stranded RNA dependent protein kinase, FAST kinase, AIM1, IPL1-like midbody-associated protein kinase-1, NIMA-like protein kinase 1 (NLK1), the cyclin-dependent kinases (cdk1-10), checkpoint kinase Chk1, Nek3 protein kinase, BMK1 beta kinase, Clk1, Clk2, Clk3, extracellular signal-regulated kinases 1, 3, and 6, cdc28 protein kinase 1, cdc28 protein kinase 2, pLK, Myt1, c-Jun N-terminal kinase 2, Cam kinase 1, the MAP kinases, insulin-stimulated protein kinase 1, beta-adrenergic receptor kinase 2, ribosomal protein S6 kinase, kinase suppressor of ras-1 (KSR1), putative serine/threonine protein kinase Prk, PkB kinase, cAMP-dependent protein kinase, cGMP-dependent protein kinase, type II cGMP-dependent protein kinase, protein kinases Dyrk2, Dyrk3, and Dyrk4, Rho-associated coiled-coil containing protein kinase p160ROCK, protein tyrosine kinase t-Ror1, Ste20-related kinases, cell adhesion kinase beta, protein kinase 3, stress-activated

protein kinase 4, protein kinase Zpk, serine kinase hPAK65, dual specificity mitogen-activated protein kinases 1 and 2, casein kinase I gamma 2, p21-activated protein kinase Pak1, lipid-activated protein kinase PRK2, focal adhesion kinase, dual-specificity tyrosine-phosphorylation regulated kinase, myosin light chain kinase, serine kinases SRPK2, TESK1, and VRK2, B lymphocyte serine/threonine protein kinase, stress-activated protein kinases JNK1 and JNK2, phosphorylase kinase, protein tyrosine kinase Tec, Jak2 kinase, protein kinase Ndr, MEK kinase 3, SHB adaptor protein (a Src homology 2 protein), agammaglobulinaemia protein-tyrosine kinase (Atk), protein kinase ATR, guanylate kinase 1, thrombopoeitin receptor and its precursors, DAG kinase epsilon, and kinases encoded by oncogenes or viral oncogenes such as v-fgr (Gardner-Rasheed), v-abl (Abelson murine leukemia viral oncogene homolog 1), v-arg (Abelson murine leukemia viral oncogene homolog, Abelson-related gene), v-fes and v-fps (feline sarcoma viral oncogene and Fujinami avian sarcoma viral oncogene homologs), proto-oncogene c-*cot*, oncogene *pim-1*, and oncogene *mas1*.

It is recognized that the compositions, kits, and methods of the invention will be of particular utility to patients having an enhanced risk of developing prostate cancer and their medical advisors. Patients recognized as having an enhanced risk of developing prostate cancer include, for example, patients having a familial history of prostate cancer, patients identified as having a mutant oncogene (*i.e.* at least one allele), and patients determined through any other established medical criteria to be at risk for cancer or other malignancy.

The level of expression of a marker in normal (*i.e.* non-cancerous) human prostate tissue can be assessed in a variety of ways. In one embodiment, this normal level of expression is assessed by assessing the level of expression of the marker in a portion of prostate cells which appears to be non-cancerous and by comparing this normal level of expression with the level of expression in a portion of the prostate cells which is suspected of being cancerous. For example, the normal level of expression of a marker may be assessed using a non-affected portion of the prostate and this normal level of expression may be compared with the level of expression of the same marker in an affected portion of the prostate. Alternately, and particularly as further information becomes available as a result of routine performance of the methods described herein, population-average values for normal expression of the markers of the invention may be

used. In other embodiments, the 'normal' level of expression of a marker may be determined by assessing expression of the marker in a patient sample obtained from a non-cancer-afflicted patient, from a patient sample obtained from a patient before the suspected onset of prostate cancer in the patient, from archived patient samples, and the like.

The invention includes compositions, kits, and methods for assessing the presence of prostate cancer cells in a sample (*e.g.* an archived tissue sample or a sample obtained from a patient). These compositions, kits, and methods are substantially the same as those described above, except that, where necessary, the compositions, kits, and methods are adapted for use with samples other than patient samples. For example, when the sample to be used is a parafinized, archived human tissue sample, it can be necessary to adjust the ratio of compounds in the compositions of the invention, in the kits of the invention, or the methods used to assess levels of marker expression in the sample. Such methods are well known in the art and within the skill of the ordinary artisan.

The invention includes a kit for assessing the presence of prostate cancer cells (*e.g.* in a sample such as a patient sample). The kit comprises a plurality of reagents, each of which is capable of binding specifically with a nucleic acid or polypeptide corresponding to a marker of the invention. Suitable reagents for binding with a polypeptide corresponding to a marker of the invention include antibodies, antibody derivatives, antibody fragments, and the like. Suitable reagents for binding with a nucleic acid (*e.g.* a genomic DNA, an mRNA, a spliced mRNA, a cDNA, or the like) include complementary nucleic acids. For example, the nucleic acid reagents may include oligonucleotides (labeled or non-labeled) fixed to a substrate, labeled oligonucleotides not bound with a substrate, pairs of PCR primers, molecular beacon probes, and the like.

The kit of the invention may optionally comprise additional components useful for performing the methods of the invention. By way of example, the kit may comprise fluids (*e.g.* SSC buffer) suitable for annealing complementary nucleic acids or for binding an antibody with a protein with which it specifically binds, one or more sample compartments, an instructional material which describes performance of a method of the

invention, a sample of normal prostate cells, a sample of prostate cancer cells, and the like.

The invention also includes a method of making an isolated hybridoma which produces an antibody useful for assessing whether a patient is afflicted with prostate cancer. In this method, a protein or protein fragment corresponding to a marker of the invention is isolated (*e.g.* by purification from a cell in which it is expressed or by transcription and translation of a nucleic acid encoding the protein *in vivo* or *in vitro* using known methods). A vertebrate, preferably a mammal such as a mouse, rat, rabbit, or sheep, is immunized using the isolated protein or protein fragment. The vertebrate may optionally (and preferably) be immunized at least one additional time with the isolated protein or protein fragment, so that the vertebrate exhibits a robust immune response to the protein or protein fragment. Splenocytes are isolated from the immunized vertebrate and fused with an immortalized cell line to form hybridomas, using any of a variety of methods well known in the art. Hybridomas formed in this manner are then screened using standard methods to identify one or more hybridomas which produce an antibody which specifically binds with the protein or protein fragment. The invention also includes hybridomas made by this method and antibodies made using such hybridomas.

The invention also includes a method of assessing the efficacy of a test compound for inhibiting prostate cancer cells. As described above, differences in the level of expression of the markers of the invention correlate with the cancerous state of prostate cells. Although it is recognized that changes in the levels of expression of certain of the markers of the invention likely result from the cancerous state of prostate cells, it is likewise recognized that changes in the levels of expression of other of the markers of the invention induce, maintain, and promote the cancerous state of those cells. Thus, compounds which inhibit prostate cancer in a patient will cause the level of expression of one or more of the markers of the invention to change to a level nearer the normal level of expression for that marker (*i.e.* the level of expression for the marker in non-cancerous prostate cells).

This method thus comprises comparing expression of a marker in a first prostate cell sample and maintained in the presence of the test compound and expression of the marker in a second prostate cell sample and maintained in the absence of the test

compound. A significant altered level of expression of a marker listed within Tables 1-9 is an indication that the test compound inhibits prostate cancer. The prostate cell samples may, for example, be aliquots of a single sample of normal prostate cells obtained from a patient, pooled samples of normal prostate cells obtained from a patient, cells of a normal prostate cell line, aliquots of a single sample of prostate cancer cells obtained from a patient, pooled samples of prostate cancer cells obtained from a patient, cells of a prostate cancer cell line, or the like. In one embodiment, the samples are prostate cancer cells obtained from a patient and a plurality of compounds known to be effective for inhibiting various prostate cancers are tested in order to identify the compound which is likely to best inhibit the prostate cancer in the patient.

This method may likewise be used to assess the efficacy of a therapy for inhibiting prostate cancer in a patient. In this method, the level of expression of one or more markers of the invention in a pair of samples (one subjected to the therapy, the other not subjected to the therapy) is assessed. As with the method of assessing the efficacy of test compounds, if the therapy induces a significant alteration in the level of expression of a marker listed within Tables 1-9 then the therapy is efficacious for inhibiting prostate cancer. As above, if samples from a selected patient are used in this method, then alternative therapies can be assessed *in vitro* in order to select a therapy most likely to be efficacious for inhibiting prostate cancer in the patient.

As described herein, prostate cancer in patients is associated with an altered level of expression of one or more markers listed within Tables 1-9. While, as discussed above, some of these changes in expression level result from occurrence of the prostate cancer, others of these changes induce, maintain, and promote the cancerous state of prostate cancer cells. Thus, prostate cancer characterized by an altered the level of expression of one or more markers listed within Tables 1-9 can be controlled or suppressed by altering expression of those markers.

Expression of a marker listed within Tables 1-9 can be inhibited in a number of ways generally known in the art. For example, an antisense oligonucleotide can be provided to the prostate cancer cells in order to inhibit transcription, translation, or both, of the marker(s). Alternately, a polynucleotide encoding an antibody, an antibody derivative, or an antibody fragment, and operably linked with an appropriate promoter/regulator region, can be provided to the cell in order to generate intracellular

antibodies which will inhibit the function or activity of the protein corresponding to the marker(s). Using the methods described herein, a variety of molecules, particularly including molecules sufficiently small that they are able to cross the cell membrane, can be screened in order to identify molecules which inhibit expression of the marker(s). The compound so identified can be provided to the patient in order to inhibit expression of the marker(s) in the prostate cancer cells of the patient.

Expression of a marker listed in within Tables 1-9 can be enhanced in a number of ways generally known in the art. For example, a polynucleotide encoding the marker and operably linked with an appropriate promoter/regulator region can be provided to prostate cancer cells of the patient in order to induce enhanced expression of the protein (and mRNA) corresponding to the marker therein. Alternatively, if the protein is capable of crossing the cell membrane, inserting itself in the cell membrane, or is normally a secreted protein, then expression of the protein can be enhanced by providing the protein (*e.g.* directly or by way of the bloodstream or another prostate-associated fluid) to prostate cancer cells in the patient.

As described above, the cancerous state of human prostate cells is correlated with changes in the levels of expression of the markers of the invention. Thus, compounds which alter expression of one or more of the markers listed in within Tables 1-9 can induce prostate cell carcinogenesis. The invention thus includes a method for assessing the human prostate cell carcinogenic potential of a test compound. This method comprises maintaining separate aliquots of human prostate cells in the presence and absence of the test compound. Expression of a marker of the invention in each of the aliquots is compared. A significant alteration in the level of expression of a marker listed within Tables 1-9 in the aliquot maintained in the presence of the test compound (relative to the aliquot maintained in the absence of the test compound) is an indication that the test compound possesses human prostate cell carcinogenic potential. The relative carcinogenic potentials of various test compounds can be assessed by comparing the degree of enhancement or inhibition of the level of expression of the relevant markers, by comparing the number of markers for which the level of expression is enhanced or inhibited, or by comparing both.

Various aspects of the invention are described in further detail in the following subsections.

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to novel isolated nucleic acid molecules that correspond to a marker of the invention, including nucleic acids which encode a polypeptide corresponding to a marker of the invention or a portion of such a polypeptide. Isolated nucleic acids of the invention also include nucleic acid molecules sufficient for use as hybridization probes to identify nucleic acid molecules that correspond to a marker of the invention, including nucleic acids which encode a polypeptide corresponding to a marker of the invention, and fragments of such nucleic acid molecules, *e.g.*, those suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Preferably, an "isolated" nucleic acid molecule is free of sequences (preferably protein-encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, 4 kB, 3 kB, 2 kB, 1 kB, 0.5 kB or 0.1 kB of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid encoding a protein corresponding to a marker listed in one or more of Tables 1-9, can be isolated using standard molecular biology techniques and the sequence information in the database records described herein. Using all or a portion of such nucleic acid sequences, nucleic acid molecules of the invention can be isolated using standard hybridization and

cloning techniques (*e.g.*, as described in Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual, 2nd ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A process for identifying a larger fragment or the full-length coding sequence of a marker of the present invention is thus also provided. Any conventional recombinant DNA techniques applicable for isolating polynucleotides may be employed. One such method involves the 5'-RACE-PCR technique, in which the poly-A mRNA that contains the coding sequence of particular interest is first reverse transcribed with a 3'-primer comprising a sequence disclosed herein. The newly synthesized cDNA strand is then tagged with an anchor primer with a known sequence, which preferably contains a convenient cloning restriction site attached at the 5'end. The tagged cDNA is then amplified with the 3'-primer (or a nested primer sharing sequence homology to the internal sequences of the coding region) and the 5'-anchor primer. The amplification may be conducted under conditions of various levels of stringency to optimize the amplification specificity. 5'-RACE-PCR can be readily performed using commercial kits (available from, *e.g.*, BRL Life Technologies Inc., Clotech) according to the manufacturer's instructions.

Isolating the complete coding sequence of a gene can also be carried out in a hybridization assay using a suitable probe. The probe preferably comprises at least 10 nucleotides, and more preferably exhibits sequence homology to the polynucleotides of the markers of the present invention. Other high throughput screens for cDNAs, such as those involving gene chip technology, can also be employed in obtaining the complete cDNA sequence.

In addition, databases exist that reduce the complexity of ESTs by assembling contiguous EST sequences into tentative genes. For example, TIGR has assembled human ESTs into a database called THC for tentative human consensus sequences. The THC database allows for a more definitive assignment compared to ESTs alone. Software programs exist (TIGR assembler and TIGEM EST assembly machine and contig assembly program (see Huang, X., 1996, *Genomes* 33:21-23)) that allow for assembling ESTs into contiguous sequences from any organism.

Alternatively, mRNA from a sample preparation is used to construct cDNA library in the ZAP Express vector following the procedure described in Velculescu *et*

al., 1997, *Science* 270:484. The ZAP Express cDNA synthesis kit (Stratagene) is used accordingly to the manufacturer's protocol. Plates containing 250 to 2000 plaques are hybridized as described in Rupert *et al.*, 1988, *Mol. Cell. Bio.* 8:3104 to oligonucleotide probes with the same conditions previously described for standard probes except that the hybridization temperature is reduced to a room temperature. Washes are performed in 6X standard-saline-citrate 0.1% SDS for 30 minutes at room temperature. The probes are labeled with ^{32}P-ATP trough use of T4 polynucleotide kinase.

A partial cDNA (3' fragment) can be isolated by 3' directed PCR reaction. This procedure is a modification of the protocol described in Polyak *et al.*, 1997, *Nature* 389:300. Briefly, the procedure uses SAGE tags in PCR reaction such that the resultant PCR product contains the SAGE tag of interest as well as additional cDNA, the length of which is defined by the position of the tag with respect to the 3' end of the cDNA. The cDNA product derived from such a transcript-driven PCR reaction can be used for many applications.

RNA from a source to express the cDNA corresponding to a given tag is first converted to double-stranded cDNA using any standard cDNA protocol. Similar conditions used to generate cDNA for SAGE library construction can be employed except that a modified oligo-dT primer is used to derive the first strand synthesis. For example, the oligonucleotide of composition 5'-B-TCC GGC GCG CCG TTT TCC CAG TCA CGA(30)-3', contains a poly-T stretch at the 3' end for hybridization and priming from poly-A tails, an M13 priming site for use in subsequent PCR steps, a 5' Biotin label (B) for capture to strepavidin-coated magnetic beads, and an AscI restriction endonuclease site for releasing the cDNA from the strepavidin-coated magnetic beads. Theoretically, any sufficiently-sized DNA region capable of hybridizing to a PCR primer can be used as well as any other 8 base pair recognizing endonuclease.

cDNA constructed utilizing this or similar modified oligo-dT primer is then processed as described in U.S. Patent No. 5,695,937 up until adapter ligation where only one adapter is ligated to the cDNA pool. After adapter ligation, the cDNA is released from the streptavidin-coated magnetic beads and is then used as a template for cDNA amplification.

Various PCR protocols can be employed using PCR priming sites within the 3' modified oligo-dT primer and the SAGE tag. The SAGE tag-derived PCR primer

employed can be of varying length dictated by 5' extension of the tag into the adaptor sequence. cDNA products are now available for a variety of applications.

This technique can be further modified by: (1) altering the length and/or content of the modified oligo-dT primer; (2) ligating adaptors other than that previously employed within the SAGE protocol; (3) performing PCR from template retained on the streptavidin-coated magnetic beads; and (4) priming first strand cDNA synthesis with non-oligo-dT based primers.

Gene trapper technology can also be used. The reagents and manufacturer's instructions for this technology are commercially available from Life Technologies, Inc., Gaithsburg, Maryland. Briefly, a complex population of single-stranded phagemid DNA containing directional cDNA inserts is enriched for the target sequence by hybridization in solution to a biotinylated oligonucleotide probe complementary to the target sequence. The hybrids are captured on streptavidin-coated paramagnetic beads. A magnet retrieves the paramagnetic beads from the solution, leaving nonhybridized single-stranded DNAs behind. Subsequently, the captured single-stranded DNA target is released from the biotinylated oligonucleotide. After release, the cDNA clone is further enriched by using a nonbiotinylated target oligonucleotide to specifically prime conversion of the single-stranded DNA. Following transformation and plating, typically 20% to 100% of the colonies represent the cDNA clone of interest. To identify the desired cDNA clone, the colonies may be screened by colony hybridization using the ^{32}P-labeled oligonucleotide, or alternatively by DNA sequencing and alignment of all sequences obtained from numerous clones to determine a consensus sequence.

A nucleic acid molecule of the invention can be amplified using cDNA, mRNA, or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule of the invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which has a nucleotide sequence complementary to the nucleotide sequence of a nucleic acid corresponding to a marker

of the invention or to the nucleotide sequence of a nucleic acid encoding a protein which corresponds to a marker of the invention. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, a nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence, wherein the full length nucleic acid sequence comprises a marker of the invention or which encodes a polypeptide corresponding to a marker of the invention. Such nucleic acids can be used, for example, as a probe or primer. The probe/primer typically is used as one or more substantially purified oligonucleotides. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, preferably about 15, more preferably about 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 or more consecutive nucleotides of a nucleic acid of the invention.

Probes based on the sequence of a nucleic acid molecule of the invention can be used to detect transcripts or genomic sequences corresponding to one or more markers of the invention. The probe comprises a label group attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as part of a diagnostic test kit for identifying cells or tissues which mis-express the protein, such as by measuring levels of a nucleic acid molecule encoding the protein in a sample of cells from a subject, *e.g.*, detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

The invention further encompasses nucleic acid molecules that differ, due to degeneracy of the genetic code, from the nucleotide sequence of nucleic acids encoding a protein which corresponds to a marker of the invention, and thus encode the same protein.

In addition to the nucleotide sequences in Tables 1-9, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequence can exist within a population (*e.g.*, the human population). Such genetic polymorphisms can exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which occur alternatively at a given genetic locus. In addition, it will be appreciated that DNA polymorphisms that

affect RNA expression levels can also exist that may affect the overall expression level of that gene (*e.g.*, by affecting regulation or degradation).

As used herein, the phrase "allelic variant" refers to a nucleotide sequence which occurs at a given locus or to a polypeptide encoded by the nucleotide sequence.

As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a polypeptide corresponding to a marker of the invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the invention.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 7, 15, 20, 25, 30, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 550, 650, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000, 3500, 4000, 4500, or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid corresponding to a marker of the invention or to a nucleic acid encoding a protein corresponding to a marker of the invention. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 75% (80%, 85%, preferably 90%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in sections 6.3.1-6.3.6 of *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989). A preferred, non-limiting example of stringent hybridization conditions for annealing two single-stranded DNA each of which is at least about 100 bases in length and/or for annealing a single-stranded DNA and a single-stranded RNA each of which is at least about 100 bases in length, are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50-65°C. Further preferred hybridization conditions are taught in Lockhart, *et al.*, Nature Biotechnology, Volume 14, 1996 August:1675-1680; Breslauer, *et al.*, Proc. Natl. Acad. Sci. USA, Volume 83, 1986 June: 3746-3750; Van Ness, *et al.*, Nucleic Acids Research,

Volume 19, No. 19, 1991 September: 5143-5151; McGraw, *et al.*, BioTechniques, Volume 8, No. 6 1990: 674-678; and Milner, *et al.*, Nature Biotechnology, Volume 15, 1997 June: 537-541, all expressly incorporated by reference.

In addition to naturally-occurring allelic variants of a nucleic acid molecule of the invention that can exist in the population, the skilled artisan will further appreciate that sequence changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein encoded thereby. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved among homologs of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologs of various species (*e.g.*, murine and human) may be essential for activity and thus would not be likely targets for alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a polypeptide of the invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from the naturally-occurring proteins which correspond to the markers of the invention, yet retain biological activity. In one embodiment, such a protein has an amino acid sequence that is at least about 40% identical, 50%, 60%, 70%, 80%, 90%, 95%, or 98% identical to the amino acid sequence of one of the proteins which correspond to the markers of the invention.

An isolated nucleic acid molecule encoding a variant protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of nucleic acids of the invention, such that one or more amino acid residue substitutions, additions, or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

The present invention encompasses antisense nucleic acid molecules, *i.e.*, molecules which are complementary to a sense nucleic acid of the invention, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule corresponding to a marker of the invention or complementary to an mRNA sequence corresponding to a marker of the invention. Accordingly, an antisense nucleic acid of the invention can hydrogen bond to (*i.e.* anneal with) a sense nucleic acid of the invention. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, *e.g.*, all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can also be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a polypeptide of the invention. The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the

antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been sub-cloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a polypeptide corresponding to a selected marker of the invention to thereby inhibit expression of the marker, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. Examples of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site or infusion of the antisense nucleic acid into a prostate-associated body fluid. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens

expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual α-units, the strands run parallel to each other (Gaultier *et al.*, 1987, *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, 1987, *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, *FEBS Lett.* 215:327-330).

The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach, 1988, *Nature* 334:585-591) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having specificity for a nucleic acid molecule encoding a polypeptide corresponding to a marker of the invention can be designed based upon the nucleotide sequence of a cDNA corresponding to the marker. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved (see Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742). Alternatively, an mRNA encoding a polypeptide of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see, *e.g.*, Bartel and Szostak, 1993, *Science* 261:1411-1418).

The invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a polypeptide of the invention can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the polypeptide (*e.g.*, the promoter and/or enhancer) to form triple helical

structures that prevent transcription of the gene in target cells. See generally Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

In various embodiments, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.*, 1996, *Bioorganic & Medicinal Chemistry* 4(1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996), *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup, 1996, *supra*; Perry-O'Keefe *et al.*, 1996, *Proc. Natl. Acad. Sci. USA* 93:14670-675).

In another embodiment, PNAs can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated which can combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNASE H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms

of base stacking, number of bonds between the nucleobases, and orientation (Hyrup, 1996, *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra*, and Finn *et al.* (1996) *Nucleic Acids Res.* 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between the PNA and the 5' end of DNA (Mag *et al.*, 1989, *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a step-wise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.*, 1996, *Nucleic Acids Res.* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser *et al.*, 1975, *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide can include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*); or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon, 1988, *Pharm. Res.* 5:539-549). To this end, the oligonucleotide can be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The invention also includes molecular beacon nucleic acids having at least one region which is complementary to a nucleic acid of the invention, such that the molecular beacon is useful for quantitating the presence of the nucleic acid of the invention in a sample. A "molecular beacon" nucleic acid is a nucleic acid comprising a pair of complementary regions and having a fluorophore and a fluorescent quencher associated therewith. The fluorophore and quencher are associated with different portions of the nucleic acid in such an orientation that when the complementary regions are annealed with one another, fluorescence of the fluorophore is quenched by the quencher. When the complementary regions of the nucleic acid are not annealed with

one another, fluorescence of the fluorophore is quenched to a lesser degree. Molecular beacon nucleic acids are described, for example, in U.S. Patent 5,876,930.

II. Isolated Proteins and Antibodies

One aspect of the invention pertains to novel isolated proteins which correspond to individual markers of the invention, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise antibodies directed against a polypeptide corresponding to a marker of the invention. In one embodiment, the native polypeptide corresponding to a marker can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, polypeptides corresponding to a marker of the invention are produced by recombinant DNA techniques. Alternative to recombinant expression, a polypeptide corresponding to a marker of the invention can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

Biologically active portions of a polypeptide corresponding to a marker of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the protein corresponding to the marker which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding protein. A biologically active portion of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the invention.

Preferred polypeptides are encoded by the nucleotide sequences of Tables 1-9. Other useful proteins are substantially identical (*e.g.*, at least about 40%, preferably 50%, 60%, 70%, 80%, 90%, 95%, or 99%) to one of these sequences and retain the functional activity of the protein of the corresponding naturally-occurring protein yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an

algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a *k*-tuple value of 2.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The invention also provides chimeric or fusion proteins corresponding to a marker of the invention. As used herein, a "chimeric protein" or "fusion protein" comprises all or part (preferably a biologically active part) of a polypeptide corresponding to a marker of the invention operably linked to a heterologous polypeptide (*i.e.*, a polypeptide other than the polypeptide corresponding to the marker). Within the fusion protein, the term "operably linked" is intended to indicate that the polypeptide of the invention and the heterologous polypeptide are fused in-frame to each

other. The heterologous polypeptide can be fused to the amino-terminus or the carboxyl-terminus of the polypeptide of the invention.

One useful fusion protein is a GST fusion protein in which a polypeptide corresponding to a marker of the invention is fused to the carboxyl terminus of GST sequences. Such fusion proteins can facilitate the purification of a recombinant polypeptide of the invention.

In another embodiment, the fusion protein contains a heterologous signal sequence at its amino terminus. For example, the native signal sequence of a polypeptide corresponding to a marker of the invention can be removed and replaced with a signal sequence from another protein. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook *et al.*, *supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

In yet another embodiment, the fusion protein is an immunoglobulin fusion protein in which all or part of a polypeptide corresponding to a marker of the invention is fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand (soluble or membrane-bound) and a protein on the surface of a cell (receptor), to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion protein can be used to affect the bioavailability of a cognate ligand of a polypeptide of the invention. Inhibition of ligand/receptor interaction can be useful therapeutically, both for treating proliferative and differentiative disorders and for modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies directed against a polypeptide of the invention in a subject, to purify ligands and in screening assays to identify molecules which inhibit the interaction of receptors with ligands.

Chimeric and fusion proteins of the invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see, *e.g.*, Ausubel *et al.*, *supra*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the invention.

A signal sequence can be used to facilitate secretion and isolation of the secreted protein or other proteins of interest. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the invention pertains to the described polypeptides having a signal sequence, as well as to polypeptides from which the signal sequence has been proteolytically cleaved (*i.e.*, the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence can be operably linked in an expression vector to a protein of interest, such as a protein which is ordinarily not secreted or is otherwise difficult to isolate. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

The present invention also pertains to variants of the polypeptides corresponding to individual markers of the invention. Such variants have an altered amino acid sequence which can function as either agonists (mimetics) or as antagonists. Variants can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally

occurring form of the protein. An antagonist of a protein can inhibit one or more of the activities of the naturally occurring form of the protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the protein of interest. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment with the naturally occurring form of the protein.

Variants of a protein of the invention which function as either agonists (mimetics) or as antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the protein of the invention for agonist or antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display). There are a variety of methods which can be used to produce libraries of potential variants of the polypeptides of the invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, 1983, *Tetrahedron* 39:3; Itakura *et al.*, 1984, *Annu. Rev. Biochem.* 53:323; Itakura *et al.*, 1984, *Science* 198:1056; Ike *et al.*, 1983 *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of the coding sequence of a polypeptide corresponding to a marker of the invention can be used to generate a variegated population of polypeptides for screening and subsequent selection of variants. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be

derived which encodes amino terminal and internal fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan, 1992, *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.*, 1993, *Protein Engineering* 6(3):327- 331).

An isolated polypeptide corresponding to a marker of the invention, or a fragment thereof, can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length polypeptide or protein can be used or, alternatively, the invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a protein of the invention comprises at least 8 (preferably 10, 15, 20, or 30 or more) amino acid residues of the amino acid sequence of one of the polypeptides of the invention, and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with a marker of the invention to which the protein corresponds. Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, *e.g.*, hydrophilic regions. Hydrophobicity sequence analysis, hydrophilicity sequence analysis, or similar analyses can be used to identify hydrophilic regions.

An immunogen typically is used to prepare antibodies by immunizing a suitable (*i.e.* immunocompetent) subject such as a rabbit, goat, mouse, or other mammal or vertebrate. An appropriate immunogenic preparation can contain, for example, recombinantly-expressed or chemically-synthesized polypeptide. The preparation can

further include an adjuvant, such as Freund's complete or incomplete adjuvant, or a similar immunostimulatory agent.

Accordingly, another aspect of the invention pertains to antibodies directed against a polypeptide of the invention. The terms "antibody" and "antibody substance" as used interchangeably herein refer to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds an antigen, such as a polypeptide of the invention, e.g., an epitope of a polypeptide of the invention. A molecule which specifically binds to a given polypeptide of the invention is a molecule which binds the polypeptide, but does not substantially bind other molecules in a sample, *e.g.*, a biological sample, which naturally contains the polypeptide. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide of the invention as an immunogen. Preferred polyclonal antibody compositions are ones that have been selected for antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred polyclonal antibody preparations are ones that contain only antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred immunogen compositions are those that contain no other human proteins such as, for example, immunogen compositions made using a non-human host cell for recombinant expression of a polypeptide of the invention. In such a manner, the only human epitope or epitopes recognized by the resulting antibody compositions raised against this immunogen will be present as part of a polypeptide or polypeptides of the invention.

The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules can be harvested or isolated from the subject (*e.g.*, from the blood or serum of the subject) and further

purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. Alternatively, antibodies specific for a protein or polypeptide of the invention can be selected or (*e.g.*, partially purified) or purified by, *e.g.*, affinity chromatography. For example, a recombinantly expressed and purified (or partially purified) protein of the invention is produced as described herein, and covalently or non-covalently coupled to a solid support such as, for example, a chromatography column. The column can then be used to affinity purify antibodies specific for the proteins of the invention from a sample containing antibodies directed against a large number of different epitopes, thereby generating a substantially purified antibody composition, *i.e.*, one that is substantially free of contaminating antibodies. By a substantially purified antibody composition is meant, in this context, that the antibody sample contains at most only 30% (by dry weight) of contaminating antibodies directed against epitopes other than those of the desired protein or polypeptide of the invention, and preferably at most 20%, yet more preferably at most 10%, and most preferably at most 5% (by dry weight) of the sample is contaminating antibodies. A purified antibody composition means that at least 99% of the antibodies in the composition are directed against the desired protein or polypeptide of the invention.

At an appropriate time after immunization, *e.g.*, when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (see Kozbor *et al.*, 1983, *Immunol. Today* 4:72), the EBV-hybridoma technique (see Cole *et al.*, pp. 77-96 In *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., 1985) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology*, Coligan *et al.* ed., John Wiley & Sons, New York, 1994). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide of interest, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody directed against a polypeptide of the invention can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an

antibody phage display library) with the polypeptide of interest. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.*, Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816,397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having one or more complementarily determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521- 3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-

449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

Antibodies of the invention may be used as therapeutic agents in treating cancers. In a preferred embodiment, completely human antibodies of the invention are used for therapeutic treatment of human cancer patients, particularly those having prostate cancer. Such antibodies can be produced, for example, using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of a polypeptide corresponding to a marker of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation.

Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, *e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, *e.g.*, a murine antibody, is used to guide the selection of a completely human antibody recognizing the same epitope (Jespers *et al.*, 1994, *Bio/technology* 12:899-903).

An antibody directed against a polypeptide corresponding to a marker of the invention (*e.g.*, a monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation.

Moreover, such an antibody can be used to detect the marker (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the level and pattern of expression of the marker. The antibodies can also be used diagnostically to monitor protein levels in tissues or body fluids (*e.g.* in an ovary-associated body fluid) as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Further, an antibody (or fragment thereof) can be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, .alpha.-interferon, .beta.-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

Accordingly, in one aspect, the invention provides substantially purified antibodies or fragments thereof, and non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences of the present invention, an amino acid sequence encoded by the cDNA of the present invention, a fragment of at least 15 amino acid residues of an amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to the amino acid sequence of the present invention (wherein the percent identity is determined using the

ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. In various embodiments, the substantially purified antibodies of the invention, or fragments thereof, can be human, non-human, chimeric and/or humanized antibodies.

In another aspect, the invention provides non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of the present invention, an amino acid sequence encoded by the cDNA of the present invention, a fragment of at least 15 amino acid residues of the amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to the amino acid sequence of the present invention (wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. Such non-human antibodies can be goat, mouse, sheep, horse, chicken, rabbit, or rat antibodies. Alternatively, the non-human antibodies of the invention can be chimeric and/or humanized antibodies. In addition, the non-human antibodies of the invention can be polyclonal antibodies or monoclonal antibodies.

In still a further aspect, the invention provides monoclonal antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences of the present invention, an amino acid sequence encoded by the cDNA of the present invention, a fragment of at least 15 amino acid residues of an amino acid sequence of the present invention, an amino acid sequence which is at least 95% identical to an amino acid sequence of the present invention (wherein the percent

identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45□C and washing in 0.2 X SSC, 0.1% SDS at 65°C. The monoclonal antibodies can be human, humanized, chimeric and/or non-human antibodies.

The substantially purified antibodies or fragments thereof may specifically bind to a signal peptide, a secreted sequence, an extracellular domain, a transmembrane or a cytoplasmic domain or cytoplasmic membrane of a polypeptide of the invention. In a particularly preferred embodiment, the substantially purified antibodies or fragments thereof, the non-human antibodies or fragments thereof, and/or the monoclonal antibodies or fragments thereof, of the invention specifically bind to a secreted sequence or an extracellular domain of the amino acid sequences of the present invention.

Any of the antibodies of the invention can be conjugated to a therapeutic moiety or to a detectable substance. Non-limiting examples of detectable substances that can be conjugated to the antibodies of the invention are an enzyme, a prosthetic group, a fluorescent material, a luminescent material, a bioluminescent material, and a radioactive material.

The invention also provides a kit containing an antibody of the invention conjugated to a detectable substance, and instructions for use. Still another aspect of the invention is a pharmaceutical composition comprising an antibody of the invention and a pharmaceutically acceptable carrier. In preferred embodiments, the pharmaceutical composition contains an antibody of the invention, a therapeutic moiety, and a pharmaceutically acceptable carrier.

Still another aspect of the invention is a method of making an antibody that specifically recognizes a polypeptide of the present invention, the method comprising immunizing a mammal with a polypeptide. The polypeptide used as an immungen comprises an amino acid sequence selected from the group consisting of the amino acid sequence of the present invention, an amino acid sequence encoded by the cDNA of the nucleic acid molecules of the present invention, a fragment of at least 15 amino acid residues of the amino acid sequence of the present invention, an amino acid sequence

which is at least 95% identical to the amino acid sequence of the present invention (wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4) and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule consisting of the nucleic acid molecules of the present invention, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C.

After immunization, a sample is collected from the mammal that contains an antibody that specifically recognizes the polypeptide. Preferably, the polypeptide is recombinantly produced using a non-human host cell. Optionally, the antibodies can be further purified from the sample using techniques well known to those of skill in the art. The method can further comprise producing a monoclonal antibody- producing cell from the cells of the mammal. Optionally, antibodies are collected from the antibody-producing cell.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a polypeptide corresponding to a marker of the invention (or a portion of such a polypeptide). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, namely expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*,

replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell. This means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Methods in Enzymology: Gene Expression Technology* vol.185, Academic Press, San Diego, CA (1991). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

The recombinant expression vectors of the invention can be designed for expression of a polypeptide corresponding to a marker of the invention in prokaryotic (*e.g.*, *E. coli*) or eukaryotic cells (*e.g.*, insect cells {using baculovirus expression vectors}, yeast cells or mammalian cells). Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either

fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988, *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, 1988, *Gene* 69:301-315) and pET 11d (Studier *et al.*, p. 60-89, In *Gene Expression Technology: Methods in Enzymology* vol.185, Academic Press, San Diego, CA, 1991). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a co-expressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, p. 119-128, In *Gene Expression Technology: Methods in Enzymology* vol. 185, Academic Press, San Diego, CA, 1990. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, 1992, *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari *et al.*, 1987, *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, 1982, *Cell* 30:933-943), pJRY88 (Schultz *et al.*, 1987, *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

Alternatively, the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.*, 1983, *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers, 1989, *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987, *Nature* 329:840) and pMT2NOPC (Kaufman *et al.*, 1987, *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook *et al.*, *supra.*

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.*, 1987, *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton, 1988, *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989, *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.*, 1983, *Cell* 33:729-740; Queen and Baltimore, 1983, *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989, *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.*, 1985, *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters

(Kessel and Gruss, 1990, *Science* 249:374-379) and the α-fetoprotein promoter (Camper and Tilghman, 1989, *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the invention. Regulatory sequences operably linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue-specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, 1986, *Trends in Genetics*, Vol. 1(1).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic (*e.g.*, *E. coli*) or eukaryotic cell (*e.g.*, insect cells, yeast or mammalian cells).

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection,

lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce a polypeptide corresponding to a marker of the invention. Accordingly, the invention further provides methods for producing a polypeptide corresponding to a marker of the invention using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a polypeptide of the invention has been introduced) in a suitable medium such that the marker is produced. In another embodiment, the method further comprises isolating the marker polypeptide from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which a sequences encoding a polypeptide corresponding to a marker of the invention have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous sequences encoding a marker protein of the invention have been introduced into their genome or homologous recombinant animals in which endogenous gene(s) encoding a polypeptide corresponding to a marker of the invention sequences have been altered. Such animals are useful for studying the function and/or activity of the polypeptide corresponding to the marker and for identifying and/or evaluating modulators of polypeptide activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the

animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a nucleic acid encoding a polypeptide corresponding to a marker of the invention into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the polypeptide of the invention to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of mRNA encoding the transgene in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying the transgene can further be bred to other transgenic animals carrying other transgenes.

To create an homologous recombinant animal, a vector is prepared which contains at least a portion of a gene encoding a polypeptide corresponding to a marker of the invention into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the gene. In a preferred embodiment, the vector

is designed such that, upon homologous recombination, the endogenous gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous protein). In the homologous recombination vector, the altered portion of the gene is flanked at its 5' and 3' ends by additional nucleic acid of the gene to allow for homologous recombination to occur between the exogenous gene carried by the vector and an endogenous gene in an embryonic stem cell. The additional flanking nucleic acid sequences are of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see, *e.g.*, Thomas and Capecchi, 1987, *Cell* 51:503 for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous gene are selected (see, *e.g.*, Li *et al.*, 1992, *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see, *e.g.*, Bradley, *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, Ed., IRL, Oxford, 1987, pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 and in PCT Publication NOS. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the

FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.*, 1991, *Science* 251:1351-1355). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813 and PCT Publication NOS. WO 97/07668 and WO 97/07669.

IV. Pharmaceutical Compositions

The nucleic acid molecules, polypeptides, and antibodies (also referred to herein as "active compounds") corresponding to a marker of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention. Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide

or nucleic acid corresponding to a marker of the invention and one or more additional active compounds.

The invention also provides methods (also referred to herein as "screening assays") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, peptoids, small molecules or other drugs) which (a) bind to the marker, or (b) have a modulatory (*e.g.*, stimulatory or inhibitory) effect on the activity of the marker or, more specifically, (c) have a modulatory effect on the interactions of the marker with one or more of its natural substrates (*e.g.*, peptide, protein, hormone, co-factor, or nucleic acid), or (d) have a modulatory effect on the expression of the marker. Such assays typically comprise a reaction between the marker and one or more assay components. The other components may be either the test compound itself, or a combination of test compound and a natural binding partner of the marker.

The test compounds of the present invention may be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. Test compounds may also be obtained by any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive; see, *e.g.*, Zuckermann *et al.*, 1994, *J. Med. Chem.* 37:2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, 1997, *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992, *Biotechniques* 13:412-421), or on beads (Lam, 1991, *Nature* 354:82-84), chips (Fodor, 1993, *Nature* 364:555-556), bacteria and/or spores, (Ladner, USP 5,223,409), plasmids (Cull *et al*, 1992, *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla *et al*, 1990, *Proc. Natl. Acad. Sci.* 87:6378-6382; Felici, 1991, *J. Mol. Biol.* 222:301-310; Ladner, *supra*.).

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a marker or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to a marker or biologically active portion thereof. Determining the ability of the test compound to directly bind to a marker can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to the marker can be determined by detecting the labeled marker compound in a complex. For example, compounds (*e.g.*, marker substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, assay components can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

In another embodiment, the invention provides assays for screening candidate or test compounds which modulate the activity of a marker or a biologically active portion thereof. In all likelihood, the marker can, *in vivo*, interact with one or more molecules, such as but not limited to, peptides, proteins, hormones, cofactors and nucleic acids. For the purposes of this discussion, such cellular and extracellular molecules are referred to herein as "binding partners" or marker "substrate".

One necessary embodiment of the invention in order to facilitate such screening is the use of the marker to identify its natural *in vivo* binding partners. There are many ways to accomplish this which are known to one skilled in the art. One example is the use of the marker protein as "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al*, 1993, *Cell* 72:223-232; Madura *et al*, 1993, *J. Biol. Chem.* 268:12046-12054; Bartel *et al* ,1993, *Biotechniques* 14:920-924;

Iwabuchi *et al*, 1993 *Oncogene* 8:1693-1696; Brent WO94/10300) in order to identify other proteins which bind to or interact with the marker (binding partners) and, therefore, are possibly involved in the natural function of the marker. Such marker binding partners are also likely to be involved in the propagation of signals by the marker or downstream elements of a marker-mediated signaling pathway. Alternatively, such marker binding partners may also be found to be inhibitors of the marker.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that encodes a marker protein fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a marker-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be readily detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the marker protein.

In a further embodiment, assays may be devised through the use of the invention for the purpose of identifying compounds which modulate (*e.g.*, affect either positively or negatively) interactions between a marker and its substrates and/or binding partners. Such compounds can include, but are not limited to, molecules such as antibodies, peptides, hormones, oligonucleotides, nucleic acids, and analogs thereof. Such compounds may also be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. The preferred assay components for use in this embodiment is an prostate cancer marker identified herein, the known binding partner and/or substrate of same, and the test compound. Test compounds can be supplied from any source.

The basic principle of the assay systems used to identify compounds that interfere with the interaction between the marker and its binding partner involves

preparing a reaction mixture containing the marker and its binding partner under conditions and for a time sufficient to allow the two products to interact and bind, thus forming a complex. In order to test an agent for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of the marker and its binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the marker and its binding partner is then detected. The formation of a complex in the control reaction, but less or no such formation in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the marker and its binding partner. Conversely, the formation of more complex in the presence of compound than in the control reaction indicates that the compound may enhance interaction of the marker and its binding partner.

The assay for compounds that interfere with the interaction of the marker with its binding partner may be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the marker or its binding partner onto a solid phase and detecting complexes anchored to the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the markers and the binding partners (*e.g.*, by competition) can be identified by conducting the reaction in the presence of the test substance, *i.e.*, by adding the test substance to the reaction mixture prior to or simultaneously with the marker and its interactive binding partner. Alternatively, test compounds that disrupt preformed complexes, *e.g.*, compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are briefly described below.

In a heterogeneous assay system, either the marker or its binding partner is anchored onto a solid surface or matrix, while the other corresponding non-anchored component may be labeled, either directly or indirectly. In practice, microtitre plates are often utilized for this approach. The anchored species can be immobilized by a number

of methods, either non-covalent or covalent, that are typically well known to one who practices the art. Non-covalent attachment can often be accomplished simply by coating the solid surface with a solution of the marker or its binding partner and drying. Alternatively, an immobilized antibody specific for the assay component to be anchored can be used for this purpose. Such surfaces can often be prepared in advance and stored.

In related embodiments, a fusion protein can be provided which adds a domain that allows one or both of the assay components to be anchored to a matrix. For example, glutathione-S-transferase/marker fusion proteins or glutathione-S-transferase/binding partner can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and either the non-adsorbed marker or its binding partner, and the mixture incubated under conditions conducive to complex formation (*e.g.*, physiological conditions). Following incubation, the beads or microtiter plate wells are washed to remove any unbound assay components, the immobilized complex assessed either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of marker binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a marker or a marker binding partner can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated marker protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In certain embodiments, the protein-immobilized surfaces can be prepared in advance and stored.

In order to conduct the assay, the corresponding partner of the immobilized assay component is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted assay components are removed (*e.g.*, by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-

immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; *e.g.*, using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which modulate (inhibit or enhance) complex formation or which disrupt preformed complexes can be detected.

In an alternate embodiment of the invention, a homogeneous assay may be used. This is typically a reaction, analogous to those mentioned above, which is conducted in a liquid phase in the presence or absence of the test compound. The formed complexes are then separated from unreacted components, and the amount of complex formed is determined. As mentioned for heterogeneous assay systems, the order of addition of reactants to the liquid phase can yield information about which test compounds modulate (inhibit or enhance) complex formation and which disrupt preformed complexes.

In such a homogeneous assay, the reaction products may be separated from unreacted assay components by any of a number of standard techniques, including but not limited to: differential centrifugation, chromatography, electrophoresis and immunoprecipitation. In differential centrifugation, complexes of molecules may be separated from uncomplexed molecules through a series of centrifugal steps, due to the different sedimentation equilibria of complexes based on their different sizes and densities (see, for example, Rivas, G., and Minton, A.P., *Trends Biochem Sci* 1993 Aug;18(8):284-7). Standard chromatographic techniques may also be utilized to separate complexed molecules from uncomplexed ones. For example, gel filtration chromatography separates molecules based on size, and through the utilization of an appropriate gel filtration resin in a column format, for example, the relatively larger complex may be separated from the relatively smaller uncomplexed components. Similarly, the relatively different charge properties of the complex as compared to the uncomplexed molecules may be exploited to differentially separate the complex from the remaining individual reactants, for example through the use of ion-exchange chromatography resins. Such resins and chromatographic techniques are well known to one skilled in the art (see, *e.g.*, Heegaard, 1998, *J Mol. Recognit.* 11:141-148; Hage and Tweed, 1997, *J. Chromatogr. B. Biomed. Sci. Appl.*, 699:499-525). Gel electrophoresis

may also be employed to separate complexed molecules from unbound species (see, *e.g.*, Ausubel *et al* (eds.), In: Current Protocols in Molecular Biology, J. Wiley & Sons, New York. 1999). In this technique, protein or nucleic acid complexes are separated based on size or charge, for example. In order to maintain the binding interaction during the electrophoretic process, nondenaturing gels in the absence of reducing agent are typically preferred, but conditions appropriate to the particular interactants will be well known to one skilled in the art. Immunoprecipitation is another common technique utilized for the isolation of a protein-protein complex from solution (see, *e.g.*, Ausubel *et al* (eds.), In: Current Protocols in Molecular Biology, J. Wiley & Sons, New York. 1999). In this technique, all proteins binding to an antibody specific to one of the binding molecules are precipitated from solution by conjugating the antibody to a polymer bead that may be readily collected by centrifugation. The bound assay components are released from the beads (through a specific proteolysis event or other technique well known in the art which will not disturb the protein-protein interaction in the complex), and a second immunoprecipitation step is performed, this time utilizing antibodies specific for the correspondingly different interacting assay component. In this manner, only formed complexes should remain attached to the beads. Variations in complex formation in both the presence and the absence of a test compound can be compared, thus offering information about the ability of the compound to modulate interactions between the marker and its binding partner.

Also within the scope of the present invention are methods for direct detection of interactions between the marker and its natural binding partner and/or a test compound in a homogeneous or heterogeneous assay system without further sample manipulation. For example, the technique of fluorescence energy transfer may be utilized (see, *e.g.*, Lakowicz *et al*, U.S. Patent No. 5,631,169; Stavrianopoulos *et al*, U.S. Patent No. 4,868,103). Generally, this technique involves the addition of a fluorophore label on a first 'donor' molecule (*e.g.*, marker or test compound) such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, 'acceptor' molecule (*e.g.*, marker or test compound), which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the

'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, spatial relationships between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter). A test substance which either enhances or hinders participation of one of the species in the preformed complex will result in the generation of a signal variant to that of background. In this way, test substances that modulate interactions between a marker and its binding partner can be identified in controlled assays.

In another embodiment, modulators of marker expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA or protein, corresponding to a marker in the cell, is determined. The level of expression of mRNA or protein in the presence of the candidate compound is compared to the level of expression of mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of marker expression based on this comparison. For example, when expression of marker mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of marker mRNA or protein expression. Conversely, when expression of marker mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of marker mRNA or protein expression. The level of marker mRNA or protein expression in the cells can be determined by methods described herein for detecting marker mRNA or protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a marker protein can be further confirmed *in vivo*, *e.g.*, in a whole animal model for cellular transformation and/or tumorigenesis.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an

agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, an marker modulating agent, an antisense marker nucleic acid molecule, an marker-specific antibody, or an marker-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

It is understood that appropriate doses of small molecule agents and protein or polypeptide agents depends upon a number of factors within the knowledge of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of these agents will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the agent to have upon the nucleic acid or polypeptide of the invention. Exemplary doses of a small molecule include milligram or microgram amounts per kilogram of subject or sample weight (*e.g.* about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). Exemplary doses of a protein or polypeptide include gram, milligram or microgram amounts per kilogram of subject or sample weight (*e.g.* about 1 microgram per kilogram to about 5 grams per kilogram, about 100 micrograms per kilogram to about 500 milligrams per kilogram, or about 1 milligram per kilogram to about 50 milligrams per kilogram). It is furthermore understood that appropriate doses of one of these agents depend upon the potency of the agent with respect to the expression or activity to be modulated. Such appropriate doses can be determined using the assays described herein. When one or more of these agents is to be administered to an animal (*e.g.* a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher can, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific agent employed, the age,

body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediamine-tetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents,

for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a polypeptide or antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium, and then incorporating the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.

Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches, and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from a pressurized container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes having monoclonal antibodies incorporated therein or thereon) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

For antibodies, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the prostate epithelium). A method for lipidation of antibodies is described by Cruikshank *et al.* (1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193.

The nucleic acid molecules corresponding to a marker of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. Patent 5,328,470), or by stereotactic injection (see, *e.g.*, Chen *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.* retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising a prostate cancer marker of the present invention is also provided. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage

media. The medium is adapted or configured for having recorded thereon a marker of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the markers of the present invention.

A variety of software programs and formats can be used to store the marker information of the present invention on the electronic apparatus readable medium. For example, the nucleic acid sequence corresponding to the markers can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of dataprocessor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the markers of the present invention.

By providing the markers of the invention in readable form, one can routinely access the marker sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the present invention in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has prostate cancer or a pre-disposition to prostate cancer, wherein the method comprises the steps of determining

the presence or absence of a prostate cancer marker and based on the presence or absence of the prostate cancer marker, determining whether the subject has prostate cancer or a pre-disposition to prostate cancer and/or recommending a particular treatment for the prostate cancer or pre- prostate cancer condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has prostate cancer or a pre-disposition to prostate cancer associated with a prostate cancer marker wherein the method comprises the steps of determining the presence or absence of the prostate cancer marker, and based on the presence or absence of the prostate cancer marker, determining whether the subject has prostate cancer or a pre-disposition to prostate cancer, and/or recommending a particular treatment for the prostate cancer or pre-prostate cancer condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has prostate cancer or a pre-disposition to prostate cancer associated with a prostate cancer marker, said method comprising the steps of receiving information associated with the prostate cancer marker receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the prostate cancer marker and/or prostate cancer, and based on one or more of the phenotypic information, the prostate cancer marker, and the acquired information, determining whether the subject has prostate cancer or a pre-disposition to prostate cancer. The method may further comprise the step of recommending a particular treatment for the prostate cancer or pre- prostate cancer condition.

The present invention also provides a business method for determining whether a subject has prostate cancer or a pre-disposition to prostate cancer, said method comprising the steps of receiving information associated with the prostate cancer marker, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to the prostate cancer marker and/or prostate cancer, and based on one or more of the phenotypic information, the prostate cancer marker, and the acquired information, determining whether the subject has prostate cancer or a pre-disposition to prostate cancer. The method may further

comprise the step of recommending a particular treatment for the prostate cancer or pre-prostate cancer condition.

The invention also includes an array comprising a prostate cancer marker of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of prostate cancer, progression of prostate cancer, and processes, such a cellular transformation associated with prostate cancer.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells. This provides, for

example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes that could serve as a molecular target for diagnosis or therapeutic intervention.

VI. Predictive Medicine

The present invention pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trails are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining the level of expression of polypeptides or nucleic acids corresponding to one or more markers of the invention, in order to determine whether an individual is at risk of developing prostate cancer. Such assays can be used for prognostic or predictive purposes to thereby prophylactically treat an individual prior to the onset of the cancer.

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs or other compounds administered either to inhibit prostate cancer or to treat or prevent any other disorder {*i.e.* in order to understand any prostate carcinogenic effects that such treatment may have}) on the expression or activity of a marker of the invention in clinical trials. These and other agents are described in further detail in the following sections.

A. Diagnostic Assays

An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid corresponding to a marker of the invention in a biological sample involves obtaining a biological sample (*e.g.* a prostate smear) from a test subject and contacting the biological sample with a compound or an agent capable of detecting the polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA, or cDNA). The detection methods of the invention can thus be used to detect mRNA, protein, cDNA, or genomic DNA, for example, in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of mRNA include Northern hybridizations and *in situ*

hybridizations. *In vitro* techniques for detection of a polypeptide corresponding to a marker of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, immunohistochemistry and immunofluorescence. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a polypeptide corresponding to a marker of the invention include introducing into a subject a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

A general principle of such diagnostic and prognostic assays involves preparing a sample or reaction mixture that may contain a marker, and a probe, under appropriate conditions and for a time sufficient to allow the marker and probe to interact and bind, thus forming a complex that can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways.

For example, one method to conduct such an assay would involve anchoring the marker or probe onto a solid phase support, also referred to as a substrate, and detecting target marker/probe complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, a sample from a subject, which is to be assayed for presence and/or concentration of marker, can be anchored onto a carrier or solid phase support. In another embodiment, the reverse situation is possible, in which the probe can be anchored to a solid phase and a sample from a subject can be allowed to react as an unanchored component of the assay.

There are many established methods for anchoring assay components to a solid phase. These include, without limitation, marker or probe molecules which are immobilized through conjugation of biotin and streptavidin. Such biotinylated assay components can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In certain embodiments, the surfaces with immobilized assay components can be prepared in advance and stored.

Other suitable carriers or solid phase supports for such assays include any material capable of binding the class of molecule to which the marker or probe belongs. Well-known supports or carriers include, but are not limited to, glass, polystyrene,

nylon, polypropylene, nylon, polyethylene, dextran, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite.

In order to conduct assays with the above mentioned approaches, the non-immobilized component is added to the solid phase upon which the second component is anchored. After the reaction is complete, uncomplexed components may be removed (*e.g.*, by washing) under conditions such that any complexes formed will remain immobilized upon the solid phase. The detection of marker/probe complexes anchored to the solid phase can be accomplished in a number of methods outlined herein.

In a preferred embodiment, the probe, when it is the unanchored assay component, can be labeled for the purpose of detection and readout of the assay, either directly or indirectly, with detectable labels discussed herein and which are well-known to one skilled in the art.

It is also possible to directly detect marker/probe complex formation without further manipulation or labeling of either component (marker or probe), for example by utilizing the technique of fluorescence energy transfer (see, for example, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos, *et al.*, U.S. Patent No. 4,868,103). A fluorophore label on the first, 'donor' molecule is selected such that, upon excitation with incident light of appropriate wavelength, its emitted fluorescent energy will be absorbed by a fluorescent label on a second 'acceptor' molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, spatial relationships between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

In another embodiment, determination of the ability of a probe to recognize a marker can be accomplished without labeling either assay component (probe or marker) by utilizing a technology such as real-time Biomolecular Interaction Analysis (BIA) (see, *e.g.*, Sjolander, S. and Urbaniczky, C., 1991, *Anal. Chem.* 63:2338-2345 and

Szabo *et al.*, 1995, *Curr. Opin. Struct. Biol.* 5:699-705). As used herein, "BIA" or "surface plasmon resonance" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

Alternatively, in another embodiment, analogous diagnostic and prognostic assays can be conducted with marker and probe as solutes in a liquid phase. In such an assay, the complexed marker and probe are separated from uncomplexed components by any of a number of standard techniques, including but not limited to: differential centrifugation, chromatography, electrophoresis and immunoprecipitation. In differential centrifugation, marker/probe complexes may be separated from uncomplexed assay components through a series of centrifugal steps, due to the different sedimentation equilibria of complexes based on their different sizes and densities (see, for example, Rivas, G., and Minton, A.P., 1993, *Trends Biochem Sci.* 18(8):284-7). Standard chromatographic techniques may also be utilized to separate complexed molecules from uncomplexed ones. For example, gel filtration chromatography separates molecules based on size, and through the utilization of an appropriate gel filtration resin in a column format, for example, the relatively larger complex may be separated from the relatively smaller uncomplexed components. Similarly, the relatively different charge properties of the marker/probe complex as compared to the uncomplexed components may be exploited to differentiate the complex from uncomplexed components, for example through the utilization of ion-exchange chromatography resins. Such resins and chromatographic techniques are well known to one skilled in the art (see, *e.g.*, Heegaard, N.H., 1998, *J. Mol. Recognit.* Winter 11(1-6):14___; Hage, D.S., and Tweed, S.A. *J Chromatogr B Biomed Sci Appl* 1997 Oct 10;699(1-2):499-525). Gel electrophoresis may also be employed to separate complexed assay components from unbound components (see, *e.g.*, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1987-1999). In this technique, protein or nucleic acid complexes are separated based on size or charge, for example. In order to maintain the binding interaction during the

electrophoretic process, non-denaturing gel matrix materials and conditions in the absence of reducing agent are typically preferred. Appropriate conditions to the particular assay and components thereof will be well known to one skilled in the art.

In a particular embodiment, the level of mRNA corresponding to the marker can be determined both by *in situ* and by *in vitro* formats in a biological sample using methods known in the art. The term "biological sample" is intended to include tissues, cells, biological fluids and isolates thereof, isolated from a subject, as well as tissues, cells and fluids present within a subject. Many expression detection methods use isolated RNA. For *in vitro* methods, any RNA isolation technique that does not select against the isolation of mRNA can be utilized for the purification of RNA from prostate cells (see, *e.g.*, Ausubel *et al.*, ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York 1987-1999). Additionally, large numbers of tissue samples can readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski (1989, U.S. Patent No. 4,843,155).

The isolated mRNA can be used in hybridization or amplification assays that include, but are not limited to, Southern or Northern analyses, polymerase chain reaction analyses and probe arrays. One preferred diagnostic method for the detection of mRNA levels involves contacting the isolated mRNA with a nucleic acid molecule (probe) that can hybridize to the mRNA encoded by the gene being detected. The nucleic acid probe can be, for example, a full-length cDNA, or a portion thereof, such as an oligonucleotide of at least 7, 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a marker of the present invention. Other suitable probes for use in the diagnostic assays of the invention are described herein. Hybridization of an mRNA with the probe indicates that the marker in question is being expressed.

In one format, the mRNA is immobilized on a solid surface and contacted with a probe, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative format, the probe(s) are immobilized on a solid surface and the mRNA is contacted with the probe(s), for example, in an Affymetrix gene chip array. A skilled artisan can readily

adapt known mRNA detection methods for use in detecting the level of mRNA encoded by the markers of the present invention.

An alternative method for determining the level of mRNA corresponding to a marker of the present invention in a sample involves the process of nucleic acid amplification, *e.g.*, by rtPCR (the experimental embodiment set forth in Mullis, 1987, U.S. Patent No. 4,683,202), ligase chain reaction (Barany, 1991, *Proc. Natl. Acad. Sci. USA*, 88:189-193), self sustained sequence replication (Guatelli *et al.*, 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *Bio/Technology* 6:1197), rolling circle replication (Lizardi *et al.*, U.S. Patent No. 5,854,033) or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers. As used herein, amplification primers are defined as being a pair of nucleic acid molecules that can anneal to 5' or 3' regions of a gene (plus and minus strands, respectively, or vice-versa) and contain a short region in between. In general, amplification primers are from about 10 to 30 nucleotides in length and flank a region from about 50 to 200 nucleotides in length. Under appropriate conditions and with appropriate reagents, such primers permit the amplification of a nucleic acid molecule comprising the nucleotide sequence flanked by the primers.

For *in situ* methods, mRNA does not need to be isolated from the prostate cells prior to detection. In such methods, a cell or tissue sample is prepared/processed using known histological methods. The sample is then immobilized on a support, typically a glass slide, and then contacted with a probe that can hybridize to mRNA that encodes the marker.

As an alternative to making determinations based on the absolute expression level of the marker, determinations may be based on the normalized expression level of the marker. Expression levels are normalized by correcting the absolute expression level of a marker by comparing its expression to the expression of a gene that is not a marker, *e.g.*, a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene, or epithelial cell-specific genes. This normalization allows the comparison of the expression level in one

sample, *e.g.*, a patient sample, to another sample, *e.g.*, a non-prostate cancer sample, or between samples from different sources.

Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a marker, the level of expression of the marker is determined for 10 or more samples of normal versus cancer cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes assayed in the larger number of samples is determined and this is used as a baseline expression level for the marker. The expression level of the marker determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that marker. This provides a relative expression level.

Preferably, the samples used in the baseline determination will be from prostate cancer or from non-prostate cancer cells of prostate tissue. The choice of the cell source is dependent on the use of the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether the marker assayed is prostate specific (versus normal cells). In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from prostate cells provides a means for grading the severity of the prostate cancer state.

In another embodiment of the present invention, a polypeptide corresponding to a marker is detected. A preferred agent for detecting a polypeptide of the invention is an antibody capable of binding to a polypeptide corresponding to a marker of the invention, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

Proteins from prostate cells can be isolated using techniques that are well known to those of skill in the art. The protein isolation methods employed can, for example, be such as those described in Harlow and Lane (Harlow and Lane, 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

A variety of formats can be employed to determine whether a sample contains a protein that binds to a given antibody. Examples of such formats include, but are not limited to, enzyme immunoassay (EIA), radioimmunoassay (RIA), Western blot analysis, immunohistochemistry and enzyme linked immunoabsorbant assay (ELISA). A skilled artisan can readily adapt known protein/antibody detection methods for use in determining whether prostate cells express a marker of the present invention.

In one format, antibodies, or antibody fragments, can be used in methods such as Western blots, immunohistochemistry or immunofluorescence techniques to detect the expressed proteins. In such uses, it is generally preferable to immobilize either the antibody, proteins, or cells containing proteins, on a solid support. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite.

One skilled in the art will know many other suitable carriers for binding antibody or antigen, and will be able to adapt such support for use with the present invention. For example, protein isolated from prostate cells can be run on a polyacrylamide gel electrophoresis and immobilized onto a solid phase support such as nitrocellulose. The support can then be washed with suitable buffers followed by treatment with the detectably labeled antibody. The solid phase support can then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on the solid support can then be detected by conventional means.

The invention also encompasses kits for detecting the presence of a polypeptide or nucleic acid corresponding to a marker of the invention in a biological sample. Such kits can be used to determine if a subject is suffering from or is at increased risk of developing prostate cancer. For example, the kit can comprise a labeled compound or agent capable of detecting a polypeptide or an mRNA encoding a polypeptide corresponding to a marker of the invention in a biological sample and means for

determining the amount of the polypeptide or mRNA in the sample (*e.g.*, an antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits can also include instructions for interpreting the results obtained using the kit.

For antibody-based kits, the kit can comprise, for example: (1) a first antibody (*e.g.*, attached to a solid support) which binds to a polypeptide corresponding to a marker of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable label.

For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, *e.g.*, a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide corresponding to a marker of the invention or (2) a pair of primers useful for amplifying a nucleic acid molecule corresponding to a marker of the invention. The kit can also comprise, *e.g.*, a buffering agent, a preservative, or a protein stabilizing agent. The kit can further comprise components necessary for detecting the detectable label (*e.g.*, an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample. Each component of the kit can be enclosed within an individual container and all of the various containers can be within a single package, along with instructions for interpreting the results of the assays performed using the kit.

B. Pharmacogenomics

Agents or modulators which have a stimulatory or inhibitory effect on expression of a marker of the invention can be administered to individuals to treat (prophylactically or therapeutically) prostate cancer in the patient. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens.

Accordingly, the level of expression of a marker of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, *e.g.*, Linder (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as "altered drug action." Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as "altered drug metabolism". These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the level of expression of a marker of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of expression of a marker of the invention.

This invention also provides a process for preparing a database comprising at least one of the markers set forth in Tables 1-9. For example, th9 polynucleotide sequences are stored in a digital storage medium such that a data processing system for standardized representation of the genes that identify a prostate cancer cell is compiled. The data processing system is useful to analyze gene expression between two cells by first selecting a cell suspected of being of a neoplastic phenotype or genotype and then isolating polynucleotides from the cell. The isolated polynucleotides are sequenced. The sequences from the sample are compared with the sequence(s) present in the database using homology search techniques. Greater than 90%, more preferably greater than 95% and more preferably, greater than or equal to 97% sequence identity between the test sequence and the polynucleotides of the present invention is a positive indication that the polynucleotide has been isolated from a prostate cancer cell as defined above.

In an alternative embodiment, the polynucleotides of this invention are sequenced and the information regarding sequence and in some embodiments, relative expression, is stored in any functionally relevant program, *e.g.*, in Compare Report using the SAGE software (available though Dr. Ken Kinzler at John Hopkins University). The Compare Report provides a tabulation of the polynucleotide sequences and their abundance for the samples normalized to a defined number of polynucleotides per library (say 25,000). This is then imported into MS-ACCESS either directly or via copying the data into an Excel spreadsheet first and then from there into MS-ACCESS for additional manipulations. Other programs such as SYBASE or Oracle that permit the comparison of polynucleotide numbers could be used as alternatives to MS-ACCESS. Enhancements to the software can be designed to incorporate these additional functions. These functions consist in standard Boolean, algebraic, and text search

operations, applied in various combinations to reduce a large input set of polynucleotides to a manageable subset of a polynucleotide of specifically defined interest.

One skilled in the art may create groups containing one or more project(s) by combining the counts of specific polynucleotides within a group (*e.g.*, GroupNormal = Normal1 + Normal2, GroupTumor1 + TumorCellLine). Additional characteristic values are also calculated for each tag in the group (*e.g.*, average count, minimum count, maximum count). One skilled in the art may calculate individual tag count ratios between groups, for example the ratio of the average GroupNormal count to the average GroupTumor count for each polynucleotide. A statistical measure of the significance of observed differences in tag counts between groups may be calculated.

C. Monitoring Clinical Trials

Monitoring the influence of agents (*e.g.*, drug compounds) on the level of expression of a marker of the invention can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent to affect marker expression can be monitored in clinical trials of subjects receiving treatment for prostate cancer. In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of one or more selected markers of the invention in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression of the marker(s) in the post-administration samples; (v) comparing the level of expression of the marker(s) in the pre-administration sample with the level of expression of the marker(s) in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent can be desirable to increase expression of the marker(s) to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent can be desirable to decrease

expression of the marker(s) to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

D. Surrogate Markers

The markers of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states, and in particular, prostate cancer. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The markers of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to

the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, antibodies may be employed in an immune-based detection system for a protein marker, or marker-specific radiolabeled probes may be used to detect a mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am, J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

The markers of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates with a specific clinical drug response or susceptibility in a subject (see, e.g., McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA or protein for specific tumor markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation in marker DNA may correlate with drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

VII. Experimental Protocol

Subtracted Libraries

Subtracted libraries are generated using a PCR based method that allows the isolation of clones expressed at higher levels in one population of mRNA (tester) compared to another population (driver). Both tester and driver mRNA populations are converted into cDNA by reverse transcription, and then PCR amplified using the SMART PCR kit from Clontech. Tester and driver cDNAs are then hybridized using the PCR-Select cDNA subtraction kit from Clontech. This technique results in both subtraction and normalization, which is an equalization of copy number of low-abundance and high-abundance sequences. After generation of the subtractive libraries, a group of 96 or more clones from each library is tested to confirm differential expression by reverse Southern hybridization.

SEQ ID NOS: 1-5846 (Tables 1 and 3) were identified through the above-described subtractive library hybridization techniques. In Tables 1 and 3, SEQ ID NOS: 1-468, 2033-2047, 2077-2309, and 3663-4013 were from Library cMhqaa; SEQ ID NOS: 469-1453, 2048-2065, 2310-3121, and 4014-4739 were from Library cMhqab; SEQ ID NOS: 1454-1755, 2066-2075, 3122-3433, and 4740-5258 were from Libraries cMhqac and cMhqag; SEQ ID NOS: 1756-1877, 2076, 3434-3662, and 5259-5572 were from Library cMhqad; SEQ ID NOS: 1878-2032 and 5573-5846 were from Library cMhqaf.

The "tester" source for two of these subtracted libraries, cMhqaa and cMhqab, was comprised of cDNA generated from stage T3NO tumors. The "driver" source for the library-designated cMhqaa was cDNA prepared from benign prostate hyperplasia and activated lymphocytes [B cells, T cells (CD4 and CD8)]. The driver cDNA for the cMhqab library was prepared from benign prostate hyperplasia cDNA. The "tester" source for the cMhqac and cMhqag subtracted libraries was stage T2 and T3 cDNA derived from prostate cancer patients with poor clinical outcome, whose cancer had recurred following surgery. The "driver" source for these subtracted libraries (cMhqac and cMhqag) was activated lymphocytes and stage T2 tumor cDNA that was obtained from patients who had a good clinical outcome and their disease had not recurred after surgery. The cMhqad subtracted library was prepared using stage T2 tumor cDNA from

patients that had a good clinical outcome as tester. The driver source for this library (cMhqad) was obtained from activated lymphocytes and stage T2 and T3 tumor samples recovered from patients whose cancer had recurred following surgery. The tester for the cMhqaf library was cDNA obtained from the prostate cancer cell lines, DU145, LNCaP and PC3. The "driver" source for the cMhqaf-subtracted library was comprised of cDNA generated from a normal prostate cell strain, PrEC.

In Tables 2 and 4, SEQ ID NOS: 1-207, 617-621, 644-899, 1482-2026, 3097, 3141-3142 and 3195 were from Library cMhqsb; SEQ ID NOS: 208- 346, 622-627, 900-1128, 2027-2461, 3098-3128, 3143-3181 and 3196-3274 were from Library cMhqsc; SEQ ID NOS: 347-468, 628-635, 1129-1310, 2462-2790, 3129-3140, 3182-3194 and 3275-3323 were from Library cMhqsd; SEQ ID NOS: 469-616, 636-643, 1311-1481, and 2791-3096 were from Library cMhqse. SEQ ID NOS: 1-3323 (Tables 2 and 4) were identified through the above-described subtractive library hybridization techniques, the "tester" source for these four subtracted libraries was comprised of cDNA generated from RNA derived from five prostate stage T2 tumors. Each subtracted library was prepared with a different driver RNA pool. The library designated as cMhqsb was prepared using RNA derived from benign prostatic hyperplasia and activated lymphocytes [B cells, T cells (CD4 & CD8)]. The library designated as cMhqsc was prepared using RNA derived from benign prostatic hyperplasia and a pool of normal tissues including Colon, Liver, Lung, Skeletal Muscle, Spleen, and White Blood Cells. The library designated as cMhqsd was prepared using RNA derived from normal prostate and a pool of normal tissues including Colon, Liver, Lung, Skeletal Muscle, Spleen, and White Blood Cells. The library designated as cMhqse was prepared using RNA derived only from benign prostatic hyperplasia tissue.

SEQ ID NOS: 5847-10246 (Table 5) were identified through the above-identified subtractive library hybridization techniques. SEQ ID NOS: 5847-6369 were from Library cMhqae. The "tester" source for Library cMhqae was comprised of cDNA generated from metastatic prostate cancer cell lines (Du145, PC3 and LNCaP). The "driver" source was cDNA prepared from PrEC, a normal prostate epithelial cell strain.

SEQ ID NOS: 6370-6736, 7793-7803, 7838-7992 and 8513-8950 were from Library cMhqah; SEQ ID NOS: 6737-7133, 7804-7816, 7993-8160 and 8951-9376 were from Library cMhqai; and SEQ ID NOS: 7134-7792, 7817-7837, 8161-8512 and 9377-

10246 were from Library cMhqak. The "tester" source for these libraries was comprised of cDNA generated from metastatic prostate cancer cell line LNCaP. This "tester" DNA was synthesized from membrane-associated polypomes mRNA separated from other mRNAs (*i.e.*, cytosolic, nuclear, etc.) by density sedimentation equilbrium. The "driver" source was cDNA prepared from the non-membrane bound fraction (free cytosolic polysomes) obtained from the same density sedimentation equilibrium experiment. This mRNA was shown to be depleted of both secreted and membrane-bound protein by profiling it on the same 20K array.

SEQ ID NOs:10247-22286 (Table 8) were identified through the above-identified subtraction library hybridization techniques. SEQ ID NOs:10247-10637, 14547-14566, 14696-14899 and 16687-17423 were from Library cMhqao. SEQ ID NOs:10638-11858, 14567-14586, 14900-15442 and 17424-18147 were from Library cMhqap. SEQ ID NOs:11859-13576, 14587-14670, 15443-16193 and 18148-18996 were from Library cMhqaq. SEQ ID NOs:13577-14546, 14671-14695, 16194-16686 and 18997-19662 were from Library cMhqar. SEQ ID NOs:19663-19808, 20576-20594, 20626-20695 and 21009-21352 were from Library cMhqaj. SEQ ID NOs:19809-20575, 20595-20625, 20696-21008 and 21353-22286 were from Library cMhqal.

The "tester" source for libraries cMhqao and cMhqap was comprised of cDNA obtained from prostate cancer lymph node metastasis RNA. The "tester" source for the library cMhqaq was comprised of cDNA obtained from prostate cancer liver metastasis RNA. The "tester" source for the library cMhqar was comprised of cDNA obtained from prostate cancer bone metastasis RNA. The "tester" source for the library cMhqaj was the same cDNA as that for the cMhqao library, but it was digested with RsaI restriction endonuclease. The "tester" source for the library cMhqal was the same cDNA as that for the cMhqap library, but it was also digested with RsaI restriction endonuclease.

The "driver" source for the cMhqao, cMhqap, cMhqaj and cMhqal libraries was comprised of cDNA obtained from good clinical outcome RNA, normal lymph node RNA and tonsil RNA, mixed in a 3: 1: 0.5 ratio, respectively. The "driver" source for the cMhqaq library was comprised of cDNA obtained from good clinical outcome RNA and normal liver RNA mixed in a 3:1 ratio. The "driver" source for the cMhqar library

was comprised of cDNA obtained from good clinical outcome RNA. By "good clinical outcome RNA" is meant RNA obtained from tumor samples that were not metastatic.

SEQ ID NOs:22287-22548 (Table 9) are novel genes found, through transcriptional profiling, to be expressed at least 2-fold or greater in the following sample comparisons:

a) normal lymph nodes, normal liver and normal prostate samples, compared to primary prostate tumor samples of good clinical outcome and liver metastasis samples (prostate cancer that has metastasized to the liver);

b) normal lymph nodes, normal liver and normal prostate samples compared to primary prostate tumor samples of good clinical outcome and bone metastasis samples (prostate cancer that has metastasized to the bone); and

c) normal lymph nodes and normal prostate samples compared to primary prostate tumor samples of good clinical outcome and lymph node samples from two different sample sources (prostate cancer that has metastasized to the lymph nodes)

These sequences are therefore indicative of metastatic prostate cancer.

Proteomics

Proteins that are secreted by normal and transformed cells in culture are analyzed to identify those proteins that are likely to be secreted by cancerous cells into body fluids. Supernatants are isolated and MWT-CO filters are used to simplify the mixture of proteins. The proteins are then digested with trypsin. The tryptic peptides are loaded onto a microcapillary HPLC column where they are separated, and eluted directly into an ion trap mass spectrometer, through a custom-made electrospray ionization source. Throughout the gradient, sequence data is acquired through fragmentation of the four most intense ions (peptides) that elute off the column, while dynamically excluding those that have already been fragmented. In this way, approximately 2000 scans worth of sequence data are obtained, corresponding to approximately 50 to 200 different proteins in the sample. These data are searched against databases using correlation analysis tools, such as MS-Tag, to identify the proteins in the supernatants.

VIII. Summary Of The Data Provided In The Tables

Table 1 shows 5846 novel nucleotide sequences and Table 2 shows 3323 novel nucleotide sequences identified through subtracted library experiments. The sequences of Tables 1 and 2 were re-interpreted and those sequences are set forth in Tables 3 and 4, respectively.

The sequences of Tables 1 and 2 were re-interpreted (as in Tables 3 and 4) and vector sequence removed and those sequences are set forth in Table 7. (See, for example, the first sequence in Table 7, identified as referring to sequence 1 from Table 1).

Table 6 shows sequences from Tables 1-4 with additional sequence. (See, for example, the first sequence in Table 6, identified as referring to sequence 2 from Table 1).

Table 5 shows 4400 novel nucleotide sequences and Table 8 shows 12040 novel nucleotide sequences identified through subtractive library experiments. Table 9 shows 262 novel nucleotide sequences.

The sequences of the present invention were determined to be novel through various BLAST searches of available databases.

In Tables 1 and 3, SEQ ID NOS: 2077-5663 are preferred, SEQ ID NOS: 2033-5580 are more preferred and SEQ ID NOS: 1-2032 are most preferred. In Tables 2 and 4, SEQ ID NOS: 644-1481 and 3141-3194 are preferred, SEQ ID NOS: 617-643 are more preferred and SEQ ID NOS: 1-616 and 3097-3140 are most preferred. In Table 5, SEQ ID NOS: 5969-6106 and 7838-8512 are preferred, SEQ ID NOS: 5959-5968 and 7793-7837 are more preferred and SEQ ID NOS: 5847-5958 and 6370-7792 are most preferred. In Table 8, SEQ ID NOs:14696-16686 and 20626-21008 are preferred, SEQ ID NOs:14547-14695 and 20576-20625 are more preferred and SEQ ID NOs:10247-14546 and 19663-20575 are most preferred.

The contents of all references, patents, published patent applications, and databases cited throughout this application are hereby incorporated by reference.

Other Embodiments

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed is:

Claims

1. An isolated nucleic acid molecule comprising a nucleotide sequ Tables 1-9, or a complement thereof.

2. A vector which contains the nucleic acid molecule of claim 1.

3. A host cell which contains the nucleic acid molecule of claim 1.

4. An isolated polypeptide which is encoded by a nucleic acid mol comprising a nucleotide sequence of Tables 1-9.

5. An antibody which selectively binds to a polypeptide of claim 4

6. A method for producing a polypeptide comprising culturing the of claim 3 under conditions in which the nucleic acid molecule is expressed.

7. A method for detecting the presence of a polypeptide of claim 4 sample comprising:

a) contacting the sample with a compound which selectivel the polypeptide; and

b) determining whether the compound binds to the polypep the sample to thereby detect the presence of a polypeptide of claim 4 in the san

8. A kit comprising a compound which selectively binds to the pol of claim 4.

9. A method for detecting the presence of a nucleic acid molecule in a sample comprising:

a) contacting the sample with a nucleic acid probe or prim selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer bi nucleic acid molecule in the sample to thereby detect the presence of a nucleic molecule of claim 1 in the sample.

10. The method of claim 9, wherein the sample comprises mRNA r and is contacted with a nucleic acid probe.

11. The method of claim 9, wherein the sample is isolated from pro tissue.

12. The method of claim 9, wherein the sample is a tumor sample.

13. A kit comprising a compound which selectively hybridizes to a acid molecule of claim 1.

14. A method of assessing whether a patient is afflicted with prosta the method comprising comparing:

a) the level of expression of a marker in a patient sample, v the marker is selected from the group consisting of the markers listed in Table:

b) the normal level of expression of the marker in a control prostate cancer sample,

wherein a significant difference between the level of expression of the : the patient sample and the normal level is an indication that the patient is afflic prostate cancer.

15. The method of claim 14, wherein the marker corresponds to a s protein.

16. The method of claim 14, wherein the marker corresponds to a tı polynucleotide or portion thereof, wherein the polynucleotide comprises the m

17. The method of claim 14, wherein the sample comprises cells ob from the patient.

18. The method of claim 17, wherein the sample is a prostate tissue

19. The method of claim 17, wherein the cells are in a fluid selectec group consisting of blood fluids, semen, prostate fluid, lymph and urine.

20. The method of claim 14, wherein the level of expression of the : the sample is assessed by detecting the presence in the sample of a protein or p fragment corresponding to the marker.

21. The method of claim 20, wherein the presence of the protein or fragment is detected using a reagent which specifically binds with the protein (fragment.

22. The method of claim 21, wherein the reagent is selected from th consisting of an antibody, an antibody derivative, and an antibody fragment.

23. The method of claim 14, wherein the level of expression of the ı the sample is assessed by detecting the presence in the sample of a transcribed polynucleotide or portion thereof, wherein the transcribed polynucleotide comı marker.

24. The method of claim 23, wherein the transcribed polynucleotide mRNA.

25. The method of claim 23, wherein the transcribed polynucleotic cDNA.

26. The method of claim 23, wherein the step of detecting further c amplifying the transcribed polynucleotide.

27. The method of claim 14, wherein the level of expression of the : the sample is assessed by detecting the presence in the sample of a transcribed polynucleotide which anneals with the marker or anneals with a portion of a polynucleotide wherein the polynucleotide comprises the marker, under string hybridization conditions.

28. The method of claim 14, wherein the level of expression of the : the sample differs from the normal level of expression of the marker in a patiei afflicted with prostate cancer by a factor of at least about 2.

29. The method of claim 14, wherein the level of expression of the : the sample differs from the normal level of expression of the marker in a patiei afflicted with prostate cancer by a factor of at least about 5.

30. The method of claim 14, comprising comparing:

a) the level of expression in the sample of each of a pluralit markers independently selected from the markers listed in Tables 1-9, and

b) the normal level of expression of each of the plurality of in samples of the same type obtained from control humans not afflicted with pı cancer,

wherein the level of expression of more than one of the markers is signi altered, relative to the corresponding normal levels of expression of the marker indication that the patient is afflicted with prostate cancer.

31. The method of claim 30, wherein the level of expression of eac markers is significantly altered, relative to the corresponding normal levels of expression of the markers, is an indication that the patient is afflicted with pro cancer.

32. The method of claim 30, wherein the plurality comprises at leas the markers.

33. The method of claim 30, wherein the plurality comprises at leas the markers.

34. A method for monitoring the progression of prostate cancer in a the method comprising:

a) detecting in a patient sample at a first point in time, the of a marker, wherein the marker is selected from the group consisting of the m listed in Tables 1-9;

b) repeating step a) at a subsequent point in time; and

c) comparing the level of expression detected in steps a) ar therefrom monitoring the progression of prostate cancer.

35. The method of claim 34, wherein the marker corresponds to a s protein.

36. The method of claim 34, wherein the marker corresponds to a tr polynucleotide or portion thereof, wherein the polynucleotide comprises the m

37. The method of claim 34, wherein the sample comprises cells ob from the patient.

38. The method of claim 37, wherein the patient sample is a prostat sample.

39. The method of claim 34, wherein between the first point in tim subsequent point in time, the patient has undergone surgery to remove prostate

40. A method of assessing the efficacy of a test compound for inhit prostate cancer in a patient, the method comprising comparing:

a) expression of a marker in a first sample obtained from t and exposed to the test compound, wherein the marker is selected from the gro consisting of the markers listed in Tables 1-9, and

b) expression of the marker in a second sample obtained fr patient, wherein the sample is not exposed to the test compound,

wherein a significantly lower level of expression of the marker in the fi sample, relative to the second sample, is an indication that the test compound i efficacious for inhibiting prostate cancer in the patient.

41. The method of claim 40, wherein the first and second samples a portions of a single sample obtained from the patient.

42. The method of claim 40, wherein the first and second samples a portions of pooled samples obtained from the patient.

43. A method of assessing the efficacy of a therapy for inhibiting p cancer in a patient, the method comprising comparing:

a) expression of a marker in the first sample obtained from patient prior to providing at least a portion of the therapy to the patient, wherei marker is selected from the group consisting of the markers listed in Tables 1-

b) expression of the marker in a second sample obtained fr patient following provision of the portion of the therapy,

wherein a significantly lower level of expression of the marker in the s sample, relative to the first sample, is an indication that the therapy is efficacio inhibiting prostate cancer in the patient.

44. A method of selecting a composition for inhibiting prostate can patient, the method comprising:

a) obtaining a sample comprising cancer cells from the pat

b) separately exposing aliquots of the sample in the presen plurality of test compositions;

c) comparing expression of a marker in each of the aliquot: the marker is selected from the group consisting of the markers listed in Table:

d) selecting one of the test compositions which alters the le expression of the marker in the aliquot containing that test composition, relativ test compositions.

45. A method of inhibiting prostate cancer in a patient, the method comprising:

a) obtaining a sample comprising cancer cells from the pati

b) separately maintaining aliquots of the sample in the pres plurality of test compositions;

c) comparing expression of a marker in each of the aliquot: the marker is selected from the group consisting of the markers listed in Table:

d) administering to the patient at least one of the test comp which alters the level of expression of the marker in the aliquot containing that composition, relative to other test compositions.

46. A kit for assessing whether a patient is afflicted with prostate ca kit comprising reagents for assessing expression of a marker selected from the consisting of the markers listed in Tables 1-9.

47. A kit for assessing the presence of prostate cancer cells, the kit comprising a nucleic acid probe wherein the probe specifically binds with a tra polynucleotide corresponding to a marker selected from the group consisting o: markers listed in Tables 1-9.

48. A kit for assessing the suitability of each of a plurality of comp inhibiting prostate cancer in a patient, the kit comprising:

a) the plurality of compounds; and

b) a reagent for assessing expression of a marker selected f group consisting of the markers listed in Tables 1-9.

49. A method of making an isolated hybridoma which produces an : useful for assessing whether a patient is afflicted with prostate cancer, the met comprising:

isolating a protein or protein fragment corresponding to selected from the group consisting of the markers listed in Tables 1-9;

immunizing a mammal using the isolated protein or prot fragment;

isolating splenocytes from the immunized mammal;

fusing the isolated splenocytes with an immortalized cel form hybridomas; and

screening individual hybridomas for production of an an which specifically binds with the protein or protein fragment to isolate the hyb

50. An antibody produced by a hybridoma made by the method of c

51. A kit for assessing the presence of human prostate cancer cells, comprising an antibody, wherein the antibody specifically binds with a protein protein fragment corresponding to a marker selected from the group consisting markers listed in Tables 1-9.

52. A method of assessing the prostate cell carcinogenic potential o compound, the method comprising:

a) maintaining separate aliquots of prostate cells in the pres absence of the test compound; and

b) comparing expression of a marker in each of the aliquo the marker is selected from the group consisting of the markers listed in Tabl

wherein a significantly altered level of expression of the marker in the ı maintained in the presence of the test compound, relative to the aliquot mainta absence of the test compound, is an indication that the test compound possesse prostate cell carcinogenic potential.

53. A kit for assessing the prostate cell carcinogenic potential of a t compound, the kit comprising prostate cells and a reagent for assessing expres: marker, wherein the marker is selected from the group consisting of the marke Tables 1-9.

54. A method of inhibiting prostate cancer in a patient at risk for de prostate cancer, the method comprising inhibiting expression of a gene corresp a marker selected from the markers listed in Tables 1-9.

55. A method of treating a patient afflicted with prostate cancer, the comprising providing to cells of the patient an antisense oligonucleotide comp to a polynucleotide corresponding to a marker selected from the markers listed 1-9.

56. A method of inhibiting prostate cancer in a patient at risk for de prostate cancer, the method comprising increasing expression of a gene corres a marker selected from the markers listed in Tables 1-9.

57. A method for determining whether prostate cancer has metastasized in a patient, the method comprising comparing:

a) the level of expression of a marker in a patient smaple, wherein the marker is selected from the group consisting of the markers listed in Tables 1-9, and

b) the normal level or non-metastatic level of expression of the marker in a control sample

wherein a significant difference betweent he level of expression in the patient sample and the normal level or non-metastatic level is an indication that the prostate cancer has mestastasized.

58. The method of claim 57, wherein the marker corresponds to a secreted protein.

59. The method of claim 57, wherein the marker corresponds to a transcribed polynucleotide or portion thereof, wherein the polynucleotide comprises the marker.

60. The method of claim 57, wherein the sample comprises cells obtained from the patient.

61. The method of claim 60, wherein the patient sample is a prostate tissue sample.

62. A method for assessing the aggressiveness or indolence of prostate cancer comprising comparing:

a) the level of expression of a marker in a sample, wherein at least one marker is selected from the markers of Tables 1-9, and

b) the normal level of expression of the marker in a control sample,

wherein a significant difference between the level of expression in the sample and the normal level is an indication that the cancer is aggressive or indolent.

63. The method of claim 62, wherein the marker corresponds to a protein.

64. The method of claim 62, wherein marker corresponds to a transcribed polynucleotide or portion thereof, wherein the polynucleotide comprises the marker.

65. The method of claim 62, wherein the sample comprises cells obtained from the patient.

66. The method of claim 65, wherein the patient sample is a prostate tissue sample.

Pages 111 to 11750 of this application contain amino acid sequence listings. They can be obtained at the address given below :

World Intellectual Property Organization
34, chemin des Colombettes

CH - 1211 GENEVE 20
Suisse

Les pages 111 à 11750 de cette demande internationale contiennent des listages de séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous :

Organisation Mondiale de la Propriété Intellectuelle
34, chemin des Colombettes

CH – 1211 GENEVE 20
Suisse

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 June 2002 (06.06.2002)

PCT

(10) International Publication Number
WO 02/44320 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/CA01/01705

(22) International Filing Date: 29 November 2001 (29.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/253,728 29 November 2000 (29.11.2000) US

(71) Applicant *(for all designated States except US)*: **XENON GENETICS INC.** [CA/CA]; #501 - 520 West 6th Ave., Vancouver, British Columbia V5Z 4H5 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **WINTHER, Michael, David** [CA/CA]; 10 Anderson Blvd., Kentville, Nova Scotia B4H 5G9 (CA). **KNICKLE, Leah, Christine** [CA/CA]; 49 Craig Dr., Kentville, Nova Scotia B4N 5H2 (CA). **HAARDT, Martin** [CA/CA]; 1404 George Street, Coldbrook, Nova Scotia B4R 1B5 (CA). **DE ANTUENO, Roberto, Justo** [CA/CA]; 1520 Ashlee Dr., Coldbrook, Nova Scotia B4R 1A1 (CA). **JENKINS, D., Kenneth** [CA/CA]; 1804 Gerlad Drive, Coldbrook, Nova Scotia B4R 1A3 (CA). **NWAKA, Solomon, O.** [CA/CA]; 2075 Langille Drive, Coldbrook, Nova Scotia B4R 1C3 (CA). **ALLEN, Stephen, John** [CA/CA]; 9049 Commercial St., Suite 150, New Minas, Nova Scotia B4N 5A4 (CA). **PONTON, Andre** [CA/CA]; 15 Oakwood Drive, Kingston, Nova Scotia B0P 1R0 (CA). **GOLDBERG, Y., Paul** [CA/CA]; 62 West 18 Ave., Vancouver, British Columbia V5Y 2A4 (CA).

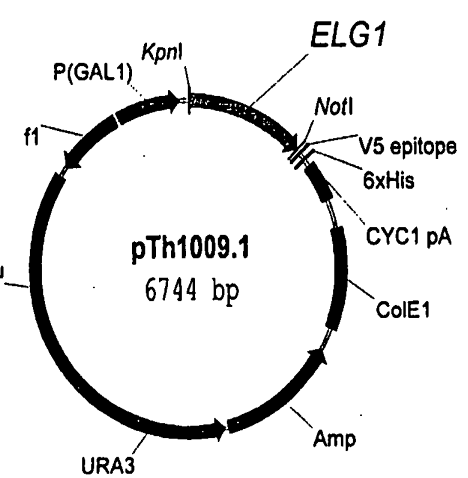
(74) Agent: **HELLER, David, J.**; Ridout & Maybee, One Queen Street East, Suite 2400, Toronto, Ontario M5C 3B1 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent

[Continued on next page]

(54) Title: HUMAN ELONGASE GENES USES THEREOF AND COMPOUNDS FOR MODULATING SAME



(57) Abstract: The present invention relates to elongase genes, their polypeptides and their control regions, and the use of such genes, polypeptides and control regions in determining compositions for use in the treatment of disease. The identified compositions regulate the expression of the elongase genes or modulate the activity of their protein products. The nucleotide and amino acid sequences are taught for ELG4, ELG6 and ELG7. The control sequences and function are taught for ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7.

WO 02/44320 A2

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Human Elongase Genes, Uses Thereof, and Compounds for Modulating Same

FIELD OF THE INVENTION

This invention relates to the identification of compounds that modulate the activity of fatty acid elongase enzymes involved in lipid metabolism and/or effectively regulate the level of expression of the elongase genes, and to compounds so identified.

BACKGROUND OF THE INVENTION

Polyunsaturated fatty acids (PUFAs) are major components of lipid compounds and complexes, such as phospholipids and lipoproteins, which provide a number of structural and functional characteristics to a wide range of biological constituents, such as the cell membranes. PUFAs are essential for the proper development, maintenance and repair of tissue. Other biological functions of PUFAs include their involvement in the expression of some genes and their role as precursor molecules for conversion into biologically active metabolites that regulate critical physiological functions. Consequently, a lack of, or imbalance in, PUFA levels has been attributed to certain pathological conditions.

Figures 1, 2 and 3 show the required desaturation and elongation steps for the production of long chain fatty acids in the n-3, n-6 and n-9/n-7 PUFA families, respectively. Fatty acid chain elongation systems have been found in liver, brain, kidney, lung, adrenals, retina, testis, small intestine and blood cells, namely leukocytes (Cinti et al., 1992, *Prog. Lipid Res.*, 31: 1-51).

Elongase genes have been identified in *Arabidopsis* (James et al., 1995, *Plant Cell*, 7: 309-319 and in *C. elegans* (WO 00/55330, Sept., 2000, Napier J.A.). Three separate elongase genes, ELO1, ELO2 and ELO3, have been identified from *S. cerevisiae*. ELO1 elongates myristic acid to palmitic acid (Toke D.A. and Martin C.E., 1996, *J. Biol. Chem.*, 271: 18413-18422) while ELO2 and ELO3 elongate long chain saturated fatty acids (Oh et al., 1997, *J. Biol. Chem.*, 272: 17376-17384).

Deficiencies in polyunsaturated fatty acids (PUFAs) have been associated with a number of diseases such as eczema, cardiovascular disorders, inflammation, psychiatric disorders, cancer, cystic fibrosis, pre-menstrual syndrome and diabetes (Horrobin D.F. [ed.], 1990, *Omega-6 Essential Fatty Acids: Pathophysiology and Roles in Clinical Medicine*, Wiley-Liss, NY and Mazza G. and Domah B.D. [eds.], 2000, *Herbs, Botanicals and Teas*, Technomic Publishers,

Lancaster, PA). Diets supplemented with PUFAs have been attempted as a treatment for a number of these conditions. The level of success for such applications has varied considerably.

Low levels of linoleic acid (18:2n-6, LA), dihomogamma-linolenic acid (20:3n-6, DGLA) and arachidonic acid (20:4n-6, AA) in adipose tissue of males have been correlated with increased mortality from coronary heart disease (Riemersma et al., 1986, *Br. Med. J. [Clin. Res. Ed.]*, 292: 1423-1427). The supplementation of LA and alpha-linolenic acid (18:3n-3, ALA) to patients suffering from hypertension did not increase the tissue levels of AA or eicosapentaenoic acid (20:5n-3, EPA) which indicates defective desaturation and elongation in the n-6 and n-3 fatty acid systems (Singer et al., 1984, *Prostaglandins Leukot. Med.*, 15: 159-165). Misoprostol, a prostaglandin E1 (PGE1) analogue, has been successfully used to treat peripheral vascular disease (Goszcz et al., 1998, *Methods Find. Exp. Clin. Pharmacol.*, 20: 439-445). PGE1 is a cyclooxygenase product of DGLA.

It has been observed that PUFAs can alleviate and correct some of the symptoms of diabetic neuropathy (Dines et al., 1993, *Diabetologia*, 36: 1132-1138 and Cotter et al., 1995, *Diabetic Neuropathy: New Concepts and Insights*, Elsevier Science B.V., Amsterdam, pp. 115-120). Researchers have speculated that the production or modulation of the cyclooxygenase and lipoxygenase metabolites of the n-3 and n-6 fatty acid families is responsible for some of these beneficial effects.

Most of the lipid metabolism disorders are characterized by a deficiency in essential fatty acids. This deficiency has been attributed to altered rate-limiting steps of delta-6-desaturation (D6D) and/or delta-5-desaturation (D5D) in PUFA biosynthesis.

SUMMARY OF INVENTION

The present invention teaches an isolated polynucleotide sequence, comprising a polynucleotide sequence which is selected from the group consisting of: (a) a sequence comprising SEQ ID NO: 4 (ELG4); (b) a sequence comprising SEQ ID NO: 8 (ELG6); (c) a sequence comprising SEQ ID NO: 11 (ELG7); (d) a sequence which is at least 80% homologous with a sequence of any of (a) to (c); (e) a sequence which is at least 90% homologous with a sequence of any of (a) to (c); (f) a sequence which is at least 95% homologous with a sequence of any of (a) to (c); (g) a sequence which is at least 98% homologous with a sequence of any of (a) to (c); (h) a sequence which is at least 99%

homologous with a sequence of any of (a) to (c); and; (i) a sequence which hybridizes to any of (a) to (h) under stringent conditions. The isolated polynucleotide sequence may be cDNA.

The invention also teaches an isolated polypeptide comprising an isolated polypeptide selected from the group consisting of: (a) a sequence comprising SEQ ID NO: 5 (ELG4); (b) a sequence comprising SEQ ID NO: 9 (ELG6); (c) a sequence comprising SEQ ID NO: 12 (ELG7); (d) a sequence which is at least 80% homologous with a sequence of any of (a) to (c); (e) a sequence which is at least 90% homologous with a sequence of any of (a) to (c); (f) a sequence which is at least 95% homologous with a sequence of any of (a) to (c); (g) a sequence which is at least 98% homologous with a sequence of any of (a) to (c); and (h) a sequence which is at least 99% homologous with a sequence of any of (a) to (c).

The invention teaches an isolated polynucleotide sequence, comprising a polynucleotide sequence which is selected from the group consisting of: (a) a sequence comprising SEQ ID NO: 1 (control region for ELG1); (b) a sequence comprising SEQ ID NO: 2 (control region for ELG2); (c) a sequence comprising SEQ ID NO: 3 (control region for ELG3); (d) a sequence comprising SEQ ID NO: 6 (control region for ELG4); (e) a sequence comprising SEQ ID NO: 7 (control region for ELG5); (f) a sequence comprising SEQ ID NO: 10 (control region for ELG6); (g) a sequence comprising SEQ ID NO: 13 (control region for ELG7); (h) a sequence which is at least 80% homologous with a sequence of any of (a) to (g); (i) a sequence which is at least 90% homologous with a sequence of any of (a) to (g); (j) a sequence which is at least 95% homologous with a sequence of any of (a) to (g); (k) a sequence which is at least 98% homologous with a sequence of any of (a) to (g); (l) a sequence which is at least 99% homologous with a sequence of any of (a) to (g); and; (m) a sequence which hybridizes to any of (a) to (l) under stringent conditions.

The invention includes an isolated polynucleotide fragment selected from the group consisting of: (a) a sequence having at least 15 sequential bases of nucleotides of a sequence of the invention; (b) a sequence having at least 30 sequential bases of nucleotides of a sequence of the invention; and (c) a sequence having at least 50 sequential bases of nucleotides of a sequence of the invention. The invention includes a polypeptide sequence which retains substantially the same biological function or activity as or is a functional derivative of a polypeptide sequence of the invention.

The invention includes an isolated polynucleotide sequence, comprising a polynucleotide sequence which retains substantially the same biological function or activity as or is a functional derivative of a polynucleotide sequence of the invention.

The invention also teaches a vector comprising a polynucleotide sequence of the invention in a suitable vector. The vector may be heterologous to the sequence. The vector may contain or encode a tag. The invention also teaches a host cell comprising a polynucleotide sequence of the invention in a host cell which is heterologous to the sequence.

The invention teaches a method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of the invention, comprising the steps of: (a) selecting a control animal having the sequence and a test animal having the sequence; (b) treating the test animal using a compound; and, (c) determining the relative quantity of an expression product of the sequence, as between the control animal and the test animal.

The invention also teaches a method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of the invention, comprising the steps of: (a) selecting a host cell of the invention; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity of an expression product of the sequence, as between the test group and the control group.

The invention further teaches a method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of the invention, comprising the steps of: (a) selecting a test group having a host cell of the invention or a part thereof, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity or relative activity of a product of the sequence or of the sequence, as between the test group and the control group.

The invention teaches a process for producing a polypeptide sequence of the invention comprising the step of culturing the host cell of the invention under conditions sufficient for the production of the polypeptide.

The invention teaches a method for identifying a compound which inhibits or promotes the activity of a polypeptide sequence of the invention, comprising the steps of: (a) selecting a control animal having the sequence and a test animal having the sequence; (b) treating the test animal using a compound; (c) determining the relative quantity or relative activity of an expression product of the sequence or of the sequence, as between the control animal and the test animal.

The invention also teaches a method for identifying a compound which inhibits or promotes the activity of a polypeptide sequence of the invention, comprising the steps of: (a) selecting a host cell of the invention; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity or relative activity of an expression product of the sequence or of the sequence, as between the test group and the control group.

The invention includes a method for identifying a compound which inhibits or promotes the activity of a polypeptide sequence of the invention, comprising the steps of: (a) selecting a test group having a host cell of the invention or a part thereof, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity or relative activity of a product of the sequence or of the sequence, as between the test group and the control group.

The invention includes a method for identifying a compound which modulates a biological activity of a polypeptide sequence of the invention, comprising the steps of: (a) providing an assay which measures a biological activity of a polypeptide sequence of the invention; (b) treating the assay with a compound; and (c) identifying a change in the biological activity of the polypeptide, wherein a difference between the treated assay and a control assay identifies the compound as modulator of the polypeptide. The polypeptide in this assay may be provided in a purified, reconstituted, cell extract or whole cell assay format, as required to assay the biological activity in question.

The invention also teaches a method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of the invention, comprising the steps of: (a) selecting a host cell of the invention; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity of an expression product of an operably linked polynucleotide to the sequence, as between the test group and the control group.

The invention also teaches a method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of the invention, comprising the steps of: (a) selecting a test group having a host cell of the invention or a part thereof, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity of an expression product of an operably linked polynucleotide to the sequence, as between the test group and the control group.

The invention includes a composition for treating a PUFA disorder comprising a compound which modulates a sequence of the invention and a pharmaceutically acceptable carrier. The invention includes the use of a composition of the invention for treating PUFA disorders.

The invention includes a method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting a germline alteration in a sequence of the invention in the subject, comprising comparing the germline sequence of a sequence of the invention from a tissue sample from the subject with the germline sequence of a wild-type of the sequence, wherein an alteration in the germline sequence of the subject indicates the presence of or a predisposition to the PUFA disorder. The invention teaches a method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject, comprising comparing the sequence of a polypeptide of the invention from a tissue sample from the subject with the sequence of a wild-type of the polypeptide, wherein an alteration in the sequence of the subject as compared to the wild-type indicates the presence of or a predisposition to the PUFA disorder.

The invention also teaches a method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a control animal having the polynucleotide and a test animal having the polynucleotide; (b) treating the test animal using a compound; and, (c) determining the relative quantity of an expression product of the polynucleotide, as between the control animal and the test animal.

The invention further teaches a method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a host cell having the polynucleotide, wherein the host cell is heterologous to the polynucleotide; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity of an expression product of the polynucleotide, as between the test group and the control group.

The invention further teaches a method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a

polynucleotide, wherein the polynucleotide is a coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a test group having a host cell with the polynucleotide or a portion of the host cell, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity or relative activity of a product of the polynucleotide or of the polynucleotide, as between the test group and the control group.

The invention teaches a method for identifying a compound modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a control animal having the polypeptide and a test animal having the polypeptide; (b) treating the test animal using a compound; (c) determining the relative quantity or relative activity of an expression product of the polypeptide or of the polypeptide, as between the control animal and the test animal.

The invention further teaches a method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a host cell comprising the polypeptide, wherein the host cell is heterologous to the polypeptide; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity or relative activity of an expression product of the polypeptide or of the polypeptide, as between the test group and the control group.

The invention also teaches a method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a test group having a host cell with the polynucleotide or a portion of the host cell, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity or relative activity of a product of the polypeptide or of the polypeptide, as between the test group and the control group.

The invention further teaches a method for identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a control region of a gene selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a control animal having the polynucleotide and a test animal having the polynucleotide; (b) treating the test animal using a compound; and, (c)

determining the relative quantity of an expression product of an operably linked polynucleotide to the polynucleotide, as between the control animal and the test animal.

The animals of the invention may be mammals. The mammals may be rats.

The invention also teaches a method for identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a control region of a gene selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a host cell comprising the polynucleotide, wherein the host cell is heterologous to the polynucleotide; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity of an expression product of an operably linked polynucleotide to the polynucleotide, as between the test group and the control group.

The invention further teaches a method for identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a control region of a gene selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a test group having a host cell with the polynucleotide or a portion of the host cell, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity of an expression product of an operably linked polynucleotide to the polynucleotide, as between the test group and the control group.

The invention includes a composition for treating a PUFA disorder comprising a compound which modulates a polynucleotide from the coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, and a pharmaceutically acceptable carrier.

The invention further teaches a composition for treating a PUFA disorder comprising a compound which modulates a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, and a pharmaceutically acceptable carrier.

The invention further teaches a composition for treating a PUFA disorder comprising a compound which modulates a polynucleotide from the control region selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, and a pharmaceutically acceptable carrier.

The compound may be selected from the group consisting of antibodies against ELG1, ELG2, ELG3 and ELG5.

The invention includes method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting a germline alteration in a polynucleotide representing the coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, from the subject, comprising comparing the germline sequence of the polynucleotide from a tissue sample from the subject with the germline sequence of a wild-type of the polynucleotide, wherein an alteration in the germline sequence of the subject indicates the presence of or a predisposition to the PUFA disorder.

The invention also teaches method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting a germline alteration in a polynucleotide representing the control region selected from the group consisting of ELG1, ELG2, ELG3 and ELG5 in the subject, comprising comparing the germline sequence of the polynucleotide from a tissue sample from the subject with the germline sequence of a wild-type of the polynucleotide, wherein an alteration in the germline sequence of the subject indicates the presence of or a predisposition to the PUFA disorder.

The invention also teaches a method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject, comprising comparing the sequence of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, from the subject with the sequence of a wild-type of the polypeptide, wherein an alteration in the sequence of the subject as compared to the wild-type indicates the presence of or a predisposition to the PUFA disorder.

The invention further teaches a method for identifying a compound which inhibits or promotes the overall activity of two or more polynucleotides, wherein the polynucleotides are control regions of two or more different genes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a host cell having the polynucleotides, wherein the host cell is heterologous to the polynucleotides;(b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantities of expression products of operably linked polynucleotides to the polynucleotides, as between the test group and the control group.

The invention further teaches a method for identifying a compound which inhibits or promotes the overall activity of two or more polynucleotides, wherein the polynucleotides are from control regions of the polynucleotides, selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) selecting a test group having a host cell with the polynucleotide or a portion of the host cell, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantities of expression products of operably linked polynucleotides to the polynucleotides, as between the test group and the control group.

The invention teaches a method for identifying a compound which inhibits or promotes the activity of two or more polynucleotides, wherein the polynucleotides are coding sequences selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence, comprising the steps of: (a) selecting a host cell having the polynucleotides, wherein the host cell are heterologous to the polynucleotides; (b) cloning the host cell and separating the clones into a test group and a control group; (c) treating the test group using a compound; and (d) determining the relative quantity or relative activity of an expression product of the polynucleotide, as between the test group and the control group.

The invention further teaches a method for identifying a compound which inhibits or promotes the activity of two or more polynucleotides, wherein the polynucleotides are coding sequences selected from the group consisting of ELG1, ELG2, ELG3; ELG4, ELG5, ELG6 and ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence, comprising the steps of: (a) selecting a test group having a host cell with the polynucleotide or a portion of the host cell, and selecting a suitable control group; (b) treating the test group using a compound; and (c) determining the relative quantity or relative activity of an expression product of the polynucleotide, as between the test group and the control group.

The invention includes a method for identifying a compound which inhibits or promotes the activity of a mammalian delta-5-desaturase enzyme and one or more enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, within the same host system, comprising the steps of: (a) providing a host system containing nucleic acid sequences which encode for a mammalian delta-5-desaturase and one or more mammalian elongase enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, operably associated with promoter regions, wherein the promoter regions are

effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence; (b) contacting the host system with a test component; (c) simultaneously evaluating the enzymatic activities of the delta-5-desaturase and the elongase enzymes, wherein a measurable difference in a level of lipid metabolites or associated cofactors in the presence of the test component compared to a control under identical conditions but in the absence of the test component is an indicator of the ability of the test component to modulate delta-5-desaturase and/or elongase enzyme activity; and (d) identifying as the compound a test component which exhibits the ability.

The invention further teaches a method for identifying a compound which inhibits or promotes the activity of a mammalian delta-6-desaturase enzyme and one or more enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, within the same host system, comprising the steps of: (a) providing a host system containing nucleic acid sequences which encode for a mammalian delta-6-desaturase and one or more mammalian elongase enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence; (b) contacting the host system with a test component; (c) simultaneously evaluating the enzymatic activities of the delta-6-desaturase and the elongase enzymes, wherein a measurable difference in a level of lipid metabolites or associated cofactors in the presence of the test component compared to a control under identical conditions but in the absence of the test component is an indicator of the ability of the test component to modulate delta-6-desaturase and/or elongase enzyme activity; and (d) identifying as the compound a test component which exhibits the ability.

The invention teaches a method for identifying a compound which inhibits or promotes the activity of a mammalian delta-5- and delta-6-desaturase enzyme and/or one or more enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, within the same host system, comprising the steps of: (a) providing a host system containing nucleic acid sequences which encode simultaneously for a mammalian delta-5-desaturase, a mammalian delta-6-desaturase and one or more mammalian elongase enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence; (b) contacting the host system with a test component; (c) simultaneously evaluating the enzymatic activities of the delta-5-desaturase, the delta-6-desaturase and the elongase enzymes, wherein a measurable difference in a level of lipid metabolites or associated cofactors in the presence of the test

component compared to a control under identical conditions but in the absence of the test component is an indicator of the ability of the test component to modulate delta-5- and/or delta-6-desaturase and/or elongase enzyme activity; and (d) identifying as the compound a test component which exhibits the ability.

The invention includes a composition for treating a PUFA disorder comprising a compound which modulates two or more human polynucleotides from control regions selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, ELG7, delta-5-desaturase, delta-6-desaturase and a pharmaceutically acceptable carrier.

The invention includes a method for detecting the presence of or the predisposition for a PUFA disorder, the method comprising determining the level of expression of two or more expression products of genes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, ELG7, delta-5-desaturase, delta-6-desaturase, in a subject relative to a predetermined control level of expression, wherein any modified expression of the expression products as compared to the control is indicative of the presence of or the predisposition for a PUFA disorder.

The invention further includes an antibody immunoreactive with a polypeptide of the invention or an immunogenic portion thereof. The invention includes an antibody immunoreactive with an elongase polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, or an immunogenic portion thereof.

The invention teaches a method for screening a medium for an elongase polypeptide of the invention or selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising: (a) labelling an antibody of the invention with a marker molecule to form a conjugate; (b) exposing the conjugate to the medium; and (c) determining whether there is binding between the conjugate and a biomolecule in the medium, wherein the binding indicates the presence of the polypeptide.

The invention teaches a method for screening a medium for an elongase polypeptide of the invention or selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising; (a) exposing an antibody of the invention to the medium; (b) exposing the antibody to a marker molecule; and (c) determining whether there is binding between the marker molecule and a biomolecule in the medium, wherein the binding indicates the presence of the polypeptide.

The invention includes compounds identified by the method of the inventions.

The invention further includes a method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting alterations in the elongation of PUFA in a peripheral blood leukocyte obtained from the subject. The invention includes a method for monitoring the development of a PUFA disorder in a subject by detecting alterations in the elongation of PUFA in a peripheral blood leukocyte obtained from the subjects. The invention further teaches a method for assessing the effect of test compounds on a PUFA disorder in a subject by assessing alterations in the elongation of PUFA in a peripheral blood leukocyte obtained from the subject.

The compounds of the invention may be selected from the group consisting of small organic molecules, peptides, polypeptides, antisense molecules, oligonucleotides, polynucleotides, fatty acids and derivatives thereof.

The invention further teaches the use of pebulate sulphoxide for the treatment of a disorder of the invention.

The disorders of the invention may be selected from the group consisting of peripheral cardiovascular disease, coronary heart disease, hypertension, atopic eczema, rheumatoid arthritis, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, psychiatric disorders, pre-menstrual syndrome, endometriosis, cystic fibrosis, alcoholism, congenital liver disease, Alzheimer's syndrome, cancer, diabetes and diabetic complications. The disorders of the invention may be selected from the group consisting of eczema, cardiovascular disorders (including but not limited to hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease, cerebrovascular disease hypertension, and peripheral vascular disease), inflammation (including but not limited to sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne), Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders (including but not limited to obesity, cachexia and anorexia), psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

BRIEF DESCRIPTION OF THE DRAWINGS

In the following description, the invention will be explained in detail with the aid of the accompanying figures, which illustrate preferred embodiments of the present invention and in which:

Figure 1 is a schematic diagram of the n-3 fatty acid metabolic pathways;

Figure 2 is a schematic diagram of the n-6 fatty acid metabolic pathways;

Figure 3 is a schematic diagram of the n-9 and n-7 fatty acid metabolic pathways;

Figure 4 is a chart showing a multiple alignment among the 7 human elongases, highlighting the invariant residues (marked by asterisks), the histidine box (marked by a box) and the ER retention signals (marked by boxes);

Figure 5 is a graph illustrating the Transmembrane Hidden Markov Model (TMHMM) prediction for transmembrane regions for ELG7;

Figure 6 is a diagram showing a topological model of a human elongase embedded in the endoplasmic reticulum;

Figure 7 is a schematic representation of plasmid pTh1009.1 (6744 bp). The human elongase (ELG1) coding sequence is shown between restriction sites for *Kpn*I and *Not*I;

Figure 8 shows the nucleotide sequence of the control region of ELG1 between position -1877 and -2865 from the translation initiation codon, ATG. This figure corresponds to SEQ. ID. NO. 1;

Figure 9 shows the nucleotide sequence of the control region of ELG2 between position -53118 and -53626 from the translation initiation codon, ATG. This figure corresponds to SEQ. ID. NO. 2;

Figure 10 shows the nucleotide sequence of the control region of ELG3 between position -37 and -1381 from the translation initiation codon, ATG. This figure corresponds to SEQ. ID. NO. 3;

Figure 11 shows the nucleotide sequence and amino acid sequence of the ELG4 gene. This figure corresponds to SEQ. ID. NOS. 4 and 5;

Figure 12 shows a 2456 bp fragment of the nucleotide sequence of the control region of ELG4. This figure corresponds to SEQ. ID. NO. 6;

Figure 13 shows the nucleotide sequence of the control region of ELG5 between position -1 and -1411 from the translation initiation codon, ATG. This figure corresponds to SEQ. ID. NO. 7;

Figure 14 shows the nucleotide sequence and amino acid sequence of the ELG6 gene. This figure corresponds to SEQ. ID. NOS. 8 and 9;

Figure 15 shows the nucleotide sequence of the control region of ELG6 between position -1 and -1937 from the translation initiation codon, ATG. This figure corresponds to SEQ. ID. NO. 10;

Figure 16 shows the nucleotide sequence and amino acid sequence of the ELG7 gene. This figure corresponds to SEQ. ID. NOS. 11 and 12;

Figure 17 shows the nucleotide sequence of the control region of ELG7 between position -1 and -2000 from the translation initiation codon, ATG. This figure corresponds to SEQ. ID. NO. 13;

Figure 18 is a schematic representation of plasmid pTh1009.2 (6743 bp). The human elongase (ELG1) coding sequence is shown between restriction sites for *Kpn*I and *Not*I;

Figure 19 is a schematic representation of plasmid pLh5015.1 (7927 bp). The human elongase (ELG3) coding sequence is shown between restriction sites for *Bam*HI and *Xba*I;

Figure 20 is a schematic representation of plasmid pGh3020.1 (6168 bp). The control region for human elongase (ELG3) is shown between two *Bgl*II restriction sites;

Figure 21 shows an HPLC analysis of radiolabelled methyl esters of fatty acids from yeast transformed with pTh1021.1 incubated with [1-^{14}C]18:3n-6, [1-^{14}C]20:4n-6, [1-^{14}C]18:3n-3 and [1-^{14}C]20:5n-3;

Figure 22 shows an HPLC analysis of radiolabelled methyl esters of fatty acids from yeast transformed with pYES2/CT incubated with [1-^{14}C]18:3n-6, [1-^{14}C]20:4n-6, [1-^{14}C]18:3n-3 and [1-^{14}C]20:5n-3;

Figure 23 shows an HPLC analysis of radiolabelled methyl esters of fatty acids from yeast co-expressing D6D/V5-His and ELG3, incubated with [1-^{14}C]18:2n-6 or [1-^{14}C]18:3n-3 and with or without galactose;

Figure 24 shows an HPLC analysis of radiolabelled methyl esters of fatty acids from yeast co-expressing D6D/V5-His and ELG3, incubated with [1-^{14}C]20:4n-6 or [1-^{14}C]20:5n-3 and with or without galactose;

Figure 25 shows an HPLC analysis of radiolabelled methyl esters of fatty acids from yeast co-expressing D5D/V5-His and ELG3, incubated with [1-^{14}C]20:4n-3 or [1-^{14}C]20:3n-6 and with or without galactose;

Figure 26 shows an HPLC analysis of radiolabelled methyl esters of fatty acids from yeast co-expressing D5D/V5-His and ELG3, incubated with [1-^{14}C]18:2n-6 or [1-^{14}C]18:3n-3 and with or without galactose.

Figure 27 shows the Northern blot analyses of ELG1, ELG2, ELG3, ELG4, ELG5 and ELG7 transcripts in a variety of human tissues.

DETAILED DESCRIPTION OF THE INVENTION

As mentioned above, research has indicated that increased levels of LA or DGLA are the result of decreased activities of delta-6 and delta-5-desaturase enzymes. The present inventors have found evidence that both the desaturase and elongase activities are affected in a PUFA related disorder.

The desaturase and elongase enzyme activities in liver microsomes from streptozotocin (STZ)-induced diabetic rats was assayed at 2 and 7 weeks post-induction. Table 1 indicates the decrease in activities compared to a control, observed during the course of the experiment. An equivalent decrease in elongation activity in STZ-induced diabetic rats has been previously reported (Suneja et al., 1990, *Biochim. Biophys. Acta*, 1042: 81-85).

Table 1

Percent Decrease of the Desaturase and Elongase Activities in Liver Microsomes from STZ-Induced Diabetic Rats

ENZYME	% Decrease	
	2 weeks	7 weeks
D6D (18:2n-6 → 18:3n-6)	28	33
Elongase (18:3n-6 → 20:3n-6)	46	43
D5D (20:3n-6 → 20:4n-6)	33	41

This data, when considered in view of what is known regarding the relationship between PUFAs and disease (above), indicates that elongase genes are involved in the development and regulation of lipid associated diseases such as inflammation, hypercholesterolemia, autoimmune disorders, atopic disorders, cystic fibrosis, psychiatric disorders, cancer, acute respiratory distress syndrome, articular cartilage degradation, arthritis, diabetes and diabetic complications. Since PUFAs are involved in a number of cell regulatory processes, the elongase genes and gene products represent realistic drug targets for the treatment or prevention of fatty acid associated diseases.

The present inventors used bioinformatic techniques to identify and analyze 7 human elongase genes (ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7). The amino acid sequences of the 7 human elongases were compared using a ClustalW algorithm (Thompson et al., 1994, *Nucl. Acids Res.*, 22: 4673-4680). One highly conserved motif, a histidine box containing 3 histidine residues, found also in a number of membrane-bound desaturases, is common to all 7 sequences. Twenty five other invariant residues, suggesting their critical importance in the catalytic activity and structure of the elongases, are identified in the multiple alignment where they are indicated by asterisk (see Figure 4).

Table 2 shows the percent identity among the 7 human elongases. The percent identities range from a low of 17% (ELG3/ELG5 and ELG3/ELG6) to a high of 55% (ELG1/ELG4).

Table 2

Percent Identities Among the 7 Human Elongases

	ELG1	ELG2	ELG3	ELG4	ELG5	ELG6	ELG7
ELG1	100						
ELG2	30	100					
ELG3	29	54	100				
ELG4	55	31	34	100			
ELG5	18	18	17	22	100		
ELG6	21	18	17	22	43	100	
ELG7	33	37	37	36	18	19	100

Based on a hidden Markov model for predicting transmembrane regions (Sonnhammer et al., 1998, In *Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology*, AAAI Press, CA, pp. 175-182), this family of seven elongases has 7 membrane spanning regions (Figure 5). These regions are highly conserved with respect to position in the amino acid sequences of the 7 elongases. The invariant histidine box is predicted to be embedded in the fourth transmembrane region. This differs from that of the membrane-bound desaturases wherein the three conserved histidine boxes are predicted to be in cytosolic loops (Shanklin et al., 1994, *Biochemistry*, 33: 12787-12794). The present model for the human elongases encompasses a ring of transmembrane domains enclosing an inner catalytic cavity for insertion of fatty acyl chains. A proposed topological model of the elongases embedded in the endoplasmic reticulum (ER) is shown in Figure 6.

The present inventors have discovered that each of the proteins has an ER retention signal (Jackson et al., 1990, *EMBO J.*, 9: 3153-3162 and Nilsson T. and Warren G., 1994, *Curr. Opin. Cell Biol.*, 6: 517-521) at the carboxyl terminus. In ER resident proteins with a type I topology (amino terminus in the lumen), the signal has been shown to consist of two critical lysines, which are in a -3 and a -4/-5 position relative to the carboxyl terminus in their cytosolic, exposed tails (K[X]KXX, where X is any amino acid). Each of the elongases has such a retention signal. Both ELG2 and ELG5, however, have modified forms of this signal

wherein the two critical lysines are found at positions -2 and -5, and -3 and -6, respectively (Figure 4).

ELG1 Gene and Polypeptide

BLASTP of the GenBank NR database with yeast ELO1 identified a protein with unknown function, CGI-88, as a potential elongase (GenBank Accession No. AAD34083). Initial cloning indicated that the cDNA sequence from which this protein was deduced (GenBank Accession No. AF151846) has a one base deletion at position 566 of the CDS. The present inventors' clone has an extra C residue at position 566 in the CDS which results in a protein, termed ELG1 by the present inventors, with a different, and longer, C-terminus than CGI-88. Since then, a gene (GenBank Accession No. AK001653) coding for a protein with no assigned function, which differs from ELG1 by one amino acid, has been submitted to GenBank (Accession No. BAA91813). The deduced amino acid sequence of ELG1 contains a F68S substitution.

The cDNA coding for ELG1 was obtained by PCR and cloned into the yeast expression vector pYES2/CT. The nucleotide sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1009.1 (Figure 7).

Yeast cells transformed with pTh1009.1 and expressing ELG1 were shown to convert 18:3n-6 to 20:3n-6, 20:4n-6 to 22:4n-6 and 24:4n-6, 18:3n-3 to 20:3n-3, and 20:5n-3 to 22:5n-3 (refer to Table 3 in Example 19). Yeast cells transformed with the pYES2/CT vector did not elongate any of these substrates. This proved that the ELG1 gene encodes a PUFA elongase. There is no published data demonstrating that this protein is a PUFA elongase. Mukerji et al. (PCT Application WO 00/12720) indicate that this protein, referred to as HS2, might be a PUFA elongase. They did not clone the coding sequence nor determine function.

The mouse ortholog of human ELG1, Ssc1 (GenBank Accession No. AF170907), has been implicated in fatty acid elongation due to its ability to complement yeast ELO mutants. Furthermore, Ssc1 gene expression correlates with elongase activity in brains of myelin-deficient mouse mutants (Tvrdik et al., 2000, *J. Cell Biol.*, 149: 707-717). Mouse Ssc1 is 92% identical and 97% similar to human ELG1.

Exons for ELG1 were mapped onto genomic DNA from human chromosome 1 (GenBank Accession No. AL139289). The gene was found to comprise 7 coding exons spanning 1.7 kb.

Using bioinformatic techniques, the control region of the ELG1 gene was identified and mapped out. By searching GenBank's EST division using BLASTN with genomic DNA and CDS for the ELG1 gene, a number of different ESTs were identified containing 5' UTR for the gene. There were 2 families of such ESTs each arising from different upstream exons which exclusively contain 5' UTR. The first exon has its 3' position at -2306 while the second exon has its 3' position at -1877 from the translation initiation codon, ATG. A 128 bp fragment of another EST (GenBank Accession No. AI373530) was also identified approximately 2.9 kb upstream of the ATG. The control region between positions -1877 and -2865 from the translation initiation codon, ATG is shown in Figure 8. A repetitive element is further identified upstream of -3600.

Northern blot studies evaluating tissue distribution showed that the ~1.3 kb ELG1 transcript is expressed in all tissues examined, with highest levels in kidney, brain, heart and placenta (Figure 27).

ELG2 Gene and Polypeptide

BLASTP of the GenBank NR database with yeast ELO1 identified a protein with unknown function (GenBank Accession No. CAB41293, since withdrawn) as a potential elongase. This protein sequence was deduced from genomic DNA (GenBank Accession No. AL034374) and represents only a partial sequence. Using GeneTrapper technology (Gibco BRL) the complete coding sequence of this protein, termed ELG2 by the present inventors, was cloned and the nucleotide sequence determined by DNA sequencing. Since then, the ELG2 coding sequence and deduced protein sequence have been submitted to GenBank (Accession Nos. AF231981 and AAF70631, respectively).

The cDNA coding for ELG2 was obtained by PCR and cloned into the yeast expression vector pYES2/CT. The sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1014.1.

Yeast cells transformed with pTh1014.1 and expressing ELG2 were shown to elongate 18:3n-6 to 20:3n-6 and 22:3n-6, 20:4n-6 to 22:4n-6 and 24:4n-6, 18:3n-3 to 20:3n-3, and 20:5n-3 to 22:5n-3 (refer to Table 3 in Example 19). Yeast transformed with the pYES2/CT vector did not elongate any of these substrates. This proved that the ELG2 gene encodes a PUFA elongase. It has been reported that this gene, referred to as HELO or HSELO, encodes a protein that is involved in the elongation of a variety of PUFAs including 18:3n-6, 20:4n-6,

18:4n-3, 20:5n-3 and 18:3n-3 (Leonard et al., 2000, *Biochem. J.*, 350: 765-770 and Mukerji et al., PCT Application WO 00/12720).

Exons for ELG2 were mapped onto genomic DNA from human chromosome 6 (GenBank Accession No. AL034374). The gene was found to comprise 7 coding exons spanning 26.5 kb.

Using bioinformatic techniques, the control region of the ELG2 gene was identified and mapped out. Using sequence data from the present inventors' clones obtained by GeneTrapper technology, 5' UTR was identified in an exon approximately 53 kb upstream of the ATG. This finding was corroborated by searching GenBank's EST division using BLASTN with the ELG2 CDS. Two ESTs were identified (GenBank Accession Nos. AA282396 and BE779576) which mapped to the same upstream exon. The control region between positions -53118 and -53626 from the translation initiation codon, ATG is shown in Figure 9. Sequence from which an EST is derived (GenBank Accession No. AA557341) lies immediately upstream of this region. A repetitive element is identified approximately 1.4 kb further upstream from the 3' end of this 5' UTR-containing exon.

Northern blot studies evaluating tissue distribution showed that the ~2.8 kb ELG2 transcript is expressed in all tissues examined, with highest levels in brain, heart and kidney, and moderate levels in the liver (Figure 27).

ELG3 Gene and Polypeptide

BLASTP of the GenBank NR database with yeast ELO1 identified a protein with unknown function (GenBank Accession No. BAA91096), as a potential elongase. This protein was deduced from cDNA (GenBank Accession No. AK000341) and is termed ELG3 by the present inventors.

The cDNA coding for ELG3 was obtained by PCR and cloned into the yeast expression vector pYES2/CT. The nucleotide sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1015.1. In comparison to GenBank Accession No. BAA91096, the protein encoded by the ELG3 gene contains two amino acid substitutions, T31M and V179I.

Yeast cells transformed with pTh1015.1 and expressing ELG3 were shown to elongate 18:3n-6 to 20:3n-6, 20:4n-6 to 22:4n-6 and 24:4n-6, 18:3n-3 to 20:3n-3, and 20:5n-3 to 22:5n-3 and

24:5n-3 (refer to Table 3 in Example 19). Yeast transformed with the pYES2/CT vector did not elongate any of these substrates. This proved that ELG3 encodes a PUFA elongase. There is no published data demonstrating that this protein is a PUFA elongase. However, Mukerji et al. (PCT Application WO 00/12720) indicate that an EST (GenBank Accession No. AI815960), found by the present inventors to represent a portion of the CDS of ELG3, may encode a partial PUFA elongase. They did not clone the coding sequence derived from this EST nor determine its function.

The mouse ortholog of human ELG3, Ssc2 (GenBank Accession No. AF170908), has been identified as putatively involved in fatty acid elongation. However, enzymatic function has not been confirmed (Tvrdik et al., 2000, *J. Cell Biol.*, 149: 707-717). Mouse Ssc2 is 88% identical and 94% similar to human ELG3.

Exons for ELG3 were mapped onto genomic DNA from human chromosome 6 (GenBank Accession No. AL121955). The gene was found to comprise 8 coding exons spanning 60.5 kb.

Using bioinformatic techniques, the control region of the ELG3 gene was identified and mapped out. By searching GenBank using BLASTN with genomic DNA and CDS for the ELG3 gene, 2 sequences (GenBank Accession Nos. BE778035 and AK000341) were identified containing 84 bp of 5' UTR immediately upstream of the initiation codon, ATG. The control region between positions -37 and -1381 from the translation initiation codon, ATG was cloned (see Example 11) and is shown in Figure 10.

Northern blot studies evaluating tissue distribution showed that the ~4.4 kb ELG3 transcript is moderately expressed in brain, with lower levels in heart, liver and placenta (Figure 27). This transcript was not detected in any of the other tissues examined.

ELG4 Gene and Polypeptide

BLASTP of the GenBank NR database with yeast ELO1 identified a protein with unknown function (GenBank Accession No. CAB70777) as a potential elongase. This protein sequence was deduced from cDNA (GenBank Accession No. AL137506) and represents only a partial sequence. Using GeneTrapper technology (Gibco BRL) and PCR amplification the full coding sequence for this protein, termed ELG4 by the present inventors, was cloned. The cDNA sequence was determined by DNA sequencing. The coding sequence and amino sequence of ELG4 are shown in Figure 11. Since then, Kawakami and coworkers have submitted a cDNA sequence to GenBank (Accession No. AK027216) that is similar to ELG4. However, in

comparison to ELG4 it does not contain the first 31 nucleotides of the coding sequence, has several nucleotide substitutions and has a one nucleotide insertion.

The cDNA coding for ELG4 was obtained by PCR and cloned into the yeast expression vector pYES2/CT. The sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1021.1.

Yeast cells transformed with pTh1021.1 and expressing ELG4 were shown to elongate 18:3n-6 to 20:3n-6 and 22:3n-6, 20:4n-6 to 22:4n-6 and 24:4n-6, and 18:3n-3 to 20:3n-3 and 22:3n-3, and 20:5n-3 to 22:5n-3 and 24:5n-3 (Refer to Table 3 in Example 19 and Figure 21). Yeast transformed with the pYES2/CT vector did not elongate any of these substrates. This proved that the ELG4 gene encodes a PUFA elongase.

Exons for ELG4 were mapped onto genomic DNA from human chromosome 5 (GenBank Accession No. AC021601). The gene was found to comprise 7 coding exons spanning at least 32 kb.

Using bioinformatic techniques, the control region of the ELG4 gene was identified and mapped out. Using sequence data from the present inventors' clones obtained by GeneTrapper technology, the 5' UTR was identified in 3 consecutive, alternatively spliced, upstream exons from the exon containing the initiation codon, ATG. The most immediate upstream exon is approximately 12 kb upstream, the next exon is over 13 kb upstream and the farthest upstream exon is at least 19 kb upstream from the ATG. The control region containing a 2456 bp fragment with its end at the 3' end of this first (most upstream) exon is shown in Figure 12. It is flanked at its 5' end by a repetitive element.

Northern blot studies evaluating tissue distribution showed that the ~4.3 kb ELG4 transcript is highly expressed in kidney and moderately expressed in brain and heart. Low levels of transcript were detected in skeletal muscle, colon, thymus, liver, small intestine and placenta (Figure 27). The transcript was not detected in spleen and peripheral blood leukocytes.

ELG5 Gene and Polypeptide

The cDNA sequence of a GenBank entry (Accession No. AK027031) encodes another potential elongase. The deduced protein sequence (GenBank Accession No. BAB15632) is termed ELG5 by the present inventors.

The cDNA coding for ELG5 was obtained by PCR and cloned into the yeast expression vector pYES2/CT. The nucleotide sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1018.1.

Yeast cells transformed with pTh1018.1 and expressing ELG5 were shown to convert 18:3n-6 to 20:3n-6 and 18:3n-3 to 20:3n-3 (refer to Table 3 in Example 19). Yeast cells transformed with the pYES2/CT vector did not elongate either of these substrates. This proved that the ELG5 gene encodes a PUFA elongase. There is no published data demonstrating that this protein is a PUFA elongase. Mukerji et al. (PCT Application WO 00/12720) indicate that HS3, which is identical to ELG5, might be a PUFA elongase. The coding sequence was cloned, however, enzymatic function was not evaluated.

Exons for ELG5 were mapped onto genomic DNA from human chromosome 4 (GenBank Accession Nos. AC004050, AC022952 and AP002080). The gene was found to comprise 4 coding exons spanning at least 88 kb.

Using bioinformatic techniques, the control region of the ELG5 gene was identified and mapped out. By searching GenBank's EST division using BLASTN with genomic DNA and CDS for the ELG5 gene, a number of different ESTs were identified containing 5' UTR for the gene. The control region between positions -1 and -1411 from the ATG is shown in Figure 13. This region is flanked at its 5' end by a repetitive element.

Northern blot studies evaluating tissue distribution showed two transcripts for ELG5 (Figure 27). The ~3.0 kb transcript is highly expressed in liver, with moderate expression in brain, colon and kidney, and low expression in heart, thymus, small intestine, placenta and skeletal muscle. The ~7.6 kb transcript is expressed in moderate levels in the brain and low levels in colon, kidney and liver.

ELG6 Gene and Polypeptide

ELG6 was identified by searching *Homo sapiens* sequences in GenBank's HTGS division with the coding sequences for ELG1, ELG2, ELG3, ELG4 and ELG5 using the TBLASTN algorithm. One sequence was identified as containing sequences similar to human elongases (GenBank Accession No. AL160011). This approach, however, failed to identify the beginning of the gene containing the translation initiation site. Therefore, further mapping and identification of ELG6 coding sequences was obtained using Cig30 (cold inducible membrane glycoprotein 30) from *Mus musculus* (GenBank Accession No. U97107), a protein found to be

similar to ELG6, as a template. The first coding exon of ELG6 containing the initiation codon, ATG, was identified in this manner.

The cDNA coding for ELG6 was obtained by PCR and cloned into the yeast expression vector pYES/CT. The nucleotide sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1041.1. The coding sequence and amino sequence of ELG6 are shown in Figure 14.

Yeast cells transformed with pTh1041.1 and expressing ELG6 were shown to elongate 18:3n-6 to 20:3n-6 and 18:3n-3 to 20:3n-3 (refer to Table 3 in Example 19). Yeast cells transformed with the pYES2/CT vector did not elongate either of these substrates. This proved that the ELG6 gene encodes a PUFA elongase.

The mouse ortholog of human ELG6, Cig30 (GenBank Accession No. U97107), has been implicated in fatty acid elongation due to its ability to complement yeast ELO2 mutants. Furthermore, Cig30 gene expression correlates with elongase activity during brown fat recruitment in mice (Tvrdik et al., 1997, *J. Biol. Chem.*, 272: 31738-31746 and Tvrdik et al., 2000, *J. Cell Biol.*, 149: 707-717). Mouse Cig30 is 69% identical and 81% similar to human ELG6.

Since the inventors' discovery of ELG6 another record has been submitted to GenBank (GenBank Accession No. AF292387) containing genomic DNA and a partial CDS for the *Homo sapiens* Cig30 ortholog. Sequence annotations, however, do not indicate the presence of the first coding exon.

Exons for ELG6 were mapped onto genomic DNA from human chromosome 10 (GenBank Accession No. AL160011). The gene was found to comprise 4 coding exons spanning approximately 2.7 kb.

Using bioinformatic techniques, the control region of the ELG6 gene was identified and mapped out. The control region between positions -1 and -1937 from the ATG is shown in Figure 15.

The transcript for ELG6 was not detected in standard Northern blot analysis in any of the tissues examined (Figure 27).

ELG7 Gene and Polypeptide

ELG7 was identified by searching *Homo sapiens* sequences in GenBank's HTGS division with the coding sequences for ELG1, ELG2, ELG3, ELG4 and ELG5 using the TBLASTN algorithm. A number of sequences were identified containing exons with sequences similar to human elongases. One such sequence,164 kb in length, (GenBank Accession No. AL132875) was found by the present inventors to contain a previously unidentified gene, termed ELG7, in 6 coding exons spanning approximately 30.5 kb of genomic DNA.

The cDNA coding for ELG7 was obtained by PCR and cloned into the yeast expression vector pYES2/CT. The nucleotide sequence was verified by DNA sequencing and the resulting plasmid was designated pTh1044.1. The coding sequence and amino sequence of ELG7 are shown in Figure 16.

Yeast cells transformed with pTh1044.1 and expressing ELG7 were shown to convert 18:3n-3 to 20:3n-3 (refer to Table 3 in Example 19). Yeast transformed with the pYES2/CT vector did not elongate this substrate. This proved that ELG7 encodes a PUFA elongase.

Using bioinformatic techniques, the control region of the ELG7 gene was identified and mapped out. By searching GenBank's EST division using BLASTN with genomic DNA for the ELG7 gene, a human EST containing 118 bp of 5' UTR for the gene was identified immediately upstream of the initiation codon, ATG (GenBank Accession No. BE878648). The control region between positions -1 and -2000 from the ATG is shown in Figure 17. A repetitive element is further identified upstream of -2700.

Northern blot studies evaluating tissue distribution showed that the ~3.0 kb ELG7 transcript is expressed in brain, thymus and placenta (Figure 27). This transcript was not detected in any of the other tissues examined.

Subject Polynucleotides and Polypeptides

The subject polynucleotides and polypeptides may be employed as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly human disease, as further discussed herein.

Nucleotide Probes

The nucleic acid molecules of the invention allow those skilled in the art to construct nucleotide probes for use in the detection of nucleotide sequences in biological materials. As described herein, a number of unique restriction sequences for restriction enzymes are incorporated in the nucleic acid molecule identified in the sequence listings of the subject polynucleotides, and these provide access to nucleotide sequences which code for polypeptides unique to the subject polynucleotides of the invention. Nucleotide sequences unique to the subject polynucleotides or isoforms thereof can also be constructed by chemical synthesis and enzymatic ligation reactions carried out by procedures known in the art.

A nucleotide probe may be labeled with a detectable marker such as a radioactive label which provides for an adequate signal and has sufficient half-life such as ^{32}P, ^{3}H, ^{14}C or the like. Other detectable markers which may be used include antigens that are recognized by a specific labeled antibody, fluorescent compounds, enzymes, antibodies specific for a labeled antigen, and chemiluminescent compounds. An appropriate label may be selected with regard to the rate of hybridization and binding of the probe to the nucleotide to be detected and the amount of nucleotide available for hybridization. The nucleotide probes may be used to detect genes related to or analogous to the subject polynucleotides of the invention.

Accordingly, the present invention also provides a method of detecting the presence of nucleic acid molecules encoding a polypeptide related to or analogous to the subject polynucleotides in a sample comprising contacting the sample under hybridization conditions with one or more of the nucleotide probes of the invention labeled with a detectable marker, and determining the degree of hybridization between the nucleic acid molecule in the sample and the nucleotide probes.

Hybridization conditions which may be used in the method of the invention are known in the art and are described for example in Sambrook et al., 1989, *Molecular Cloning, 2nd Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY. The hybridization product may be assayed using techniques known in the art. The nucleotide probe may be labeled with a detectable marker as described herein and the hybridization product may be assayed by detecting the detectable marker or the detectable change produced by the detectable marker.

Primers

The identification of the nucleic acid molecule of the invention also permits the identification

and isolation, or synthesis of nucleotide sequences which may be used as primers to amplify a polynucleotide molecule of the invention, for example in polymerase chain reaction (PCR). The length and bases of the primers for use in the PCR are selected so that they will hybridize to different strands of the desired sequence and at relative positions along the sequence such that an extension product synthesized from one primer when it is separated from its template can serve as a template for extension of the other primer into a nucleic acid of defined length.

Primers which may be used in the invention are oligonucleotides i.e. molecules containing two or more deoxyribonucleotides of the nucleic acid molecule of the invention which occur naturally as in a purified restriction endonuclease digest or are produced synthetically using techniques known in the art such as, for example, phosphotriester and phosphodiester methods or automated techniques (Connolly B. A., 1987, *Nucl. Acid Res.*, 15: 3131-3139). The primers are capable of acting as a point of initiation of synthesis when placed under conditions which permit the synthesis of a primer extension product which is complementary to the DNA sequence of the invention e.g. in the presence of nucleotide substrates, an agent for polymerization such as DNA polymerase and at suitable temperature and pH. Preferably, the primers are sequences that do not form secondary structures by base pairing with other copies of the primer or sequences that form a hair pin configuration. The primer may be single or double-stranded. When the primer is double-stranded it may be treated to separate its strands before using it to prepare amplification products. The primer preferably contains between about 7 and 25 nucleotides.

The primers may be labeled with detectable markers which allow for detection of the amplified products. Suitable detectable markers are radioactive markers such as ^{32}P, ^{35}S, ^{125}I and ^{3}H, luminescent markers such as chemiluminescent markers, preferably luminol and fluorescent markers, preferably dansyl chloride, fluorescein-5-isothiocyanate and 4-fluor-7-nitrobenz-2-oxa-1,3 diazole and cofactors such as biotin. It will be appreciated that the primers may contain non-complementary sequences provided that a sufficient amount of the primer contains a sequence which is complementary to a nucleic acid molecule of the invention or oligonucleotide sequence thereof, which is to be amplified. Restriction site linkers may also be incorporated into the primers allowing for digestion of the amplified products with the appropriate restriction enzymes facilitating cloning and sequencing of the amplified product.

Assays - Amplifying Sequences

Thus, a method of determining the presence of a nucleic acid molecule having a sequence encoding the subject polynucleotides or a predetermined oligonucleotide fragment thereof in a

sample, is provided comprising treating the sample with primers which are capable of amplifying the nucleic acid molecule or the predetermined oligonucleotide fragment thereof in a polymerase chain reaction to form amplified sequences, under conditions which permit the formation of amplified sequences and, assaying for amplified sequences.

The polymerase chain reaction refers to a process for amplifying a target nucleic acid sequence as generally described in Innis M.A. and Gelfand D.H., 1989, PCR Protocols, A Guide to Methods and Applications, Innis M.A., Gelfand D.H., Shinsky J.J. and White T.J. (eds), Academic Press, NY, pp. 3-12, which are incorporated herein by reference. Conditions for amplifying a nucleic acid template are described in Innis M.A. and Gelfand D.H., 1989, PCR Protocols, A Guide to Methods and Applications, Innis M.A., Gelfand D.H., Shinsky J.J. and White T.J. (eds), Academic Press, NY, pp. 3-12, which is also incorporated herein by reference.

The amplified products can be isolated and distinguished based on their respective sizes using techniques known in the art. For example, after amplification, the DNA sample can be separated on an agarose gel and visualized, after staining with ethidium bromide, under ultraviolet (UV) light. DNA may be amplified to a desired level and a further extension reaction may be performed to incorporate nucleotide derivatives having detectable markers such as radioactive labeled or biotin labeled nucleoside triphosphates. The primers may also be labeled with detectable markers. The detectable markers may be analyzed by restriction and electrophoretic separation or other techniques known in the art.

The conditions which may be employed in the methods of the invention using PCR are those which permit hybridization and amplification reactions to proceed in the presence of DNA in a sample and appropriate complementary hybridization primers. Conditions suitable for the polymerase chain reaction are generally known in the art. For example, see Innis M.A. and Gelfand D.H., 1989, PCR Protocols, A Guide to Methods and Applications, Innis M.A., Gelfand D.H., Shinsky J.J. and White T.J. (eds), Academic Press, NY, pp. 3-12, which is incorporated herein by reference. Preferably, the PCR utilizes polymerase obtained from thermophilic bacterium *Thermus aquaticus* (Taq polymerase, GeneAmp Kit, Perkin Elmer Cetus) or other thermostable polymerase may be used to amplify DNA template strands.

It will be appreciated that other techniques such as the Ligase Chain Reaction (LCR) and Nucleic-Acid Sequence Based Amplification (NASBA) may be used to amplify a nucleic acid molecule of the invention. In LCR, two primers which hybridize adjacent to each other on the target strand are ligated in the presence of the target strand to produce a complementary strand

(Backman, 1991 and European Published Application No. 0320308, published Jun. 14, 1989). NASBA is a continuous amplification method using two primers, one incorporating a promoter sequence recognized by an RNA polymerase and the second derived from the complementary sequence of the target sequence to the first primer (U.S. Pat. No. 5,130,238 to Malek).

Vectors

The present invention also teaches vectors which comprise a polynucleotide or polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polynucleotides of the invention by recombinant techniques.

In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, or a single or double-stranded RNA or DNA viral vector. In certain embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particular among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. All of these may be used for expression in accordance with this aspect of the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors for use in bacteria are pQE-9, pQE-16, pQE-30, pQE-40, pQE-50 and pQE-60 (Qiagen); pCRII, pCRII-TOPO, pTrcHis and pBAD-TOPO (Invitrogen); pGEM-3Z, pGEMEX-1, pET-5 (Promega); pBS phagemid vectors, Phagescript vectors, Bluescript vectors, pCAL, pET-3 and pSPUTK (Stratagene); pTrc99A, pKK223-3, pKK232-8 and pRIT2T (Pharmacia); pMAL (New England Biolabs); and pBR322 (ATCC 37017). Among eukaryotic vectors are pGAPZ, pYES2, pYES2/CT and pcDNA3.1(Invitrogen); pCAT3 and

pGL3 (Promega); pCMV-Script, pXT1, pDual, pCMVLacI, pESC, HybriZAP2.1, ImmunoZAP and pRS (Stratagene); and pSVK3, pSVL and pMSG (Pharmacia). These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example, introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide or polynucleotide in a host may be used for expression in this regard.

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase (CAT) transcription unit, downstream of restriction site or sites for introducing a candidate promoter fragment; i.e., a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the CAT gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available, such as pKK232-8 and pCAT3. Promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene. Among known prokaryotic promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli lacI* and *lacZ* promoters, the T3 and T7 promoters, the *gpt* promoter, the lambda PR and PL promoters, and the *trp* promoter. Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus (RSV), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Vectors for propagation and expression generally will include selectable markers and amplification regions, such as, for example, those set forth in Sambrook et al., *supra*.

Host Cells

As hereinbefore mentioned, the present invention also teaches host cells which are genetically engineered with vectors of the invention.

Polynucleotide constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. The subject polynucleotides or polypeptides products or isoforms or parts thereof, may be obtained by expression in a suitable host cell using techniques known in the art. Suitable host cells include prokaryotic or eukaryotic organisms or cell lines, for example bacterial, mammalian, yeast, or other fungi, viral, plant or insect cells. Methods for transforming or transfecting cells to express foreign DNA are well known in the art (See for example, Itakura et al., U.S. Pat. No. 4,704,362; Murray et al., U.S. Pat. No. 4,801,542; McKnight et al., U.S. Pat. No. 4,935,349; Hagen et al., U.S. Pat. No. 4,784,950; Axel et al., U.S. Pat. No. 4,399,216; Goeddal et al., U.S. Pat. No. 4,766,075 and Sambrook et al., 1989, *Molecular Cloning, 2nd Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY all of which are incorporated herein by reference). Representative examples of appropriate hosts include bacterial cells, such as Streptococci, Staphylococci, *E. coli*, Streptomyces and *Bacillus subtilis*; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS-1, ZR-75-1, Chang, HeLa, C127, 3T3, HepG2, BHK, 293 and Bowes melanoma cells; and plant cells.

Host cells can be genetically engineered to incorporate polynucleotides and express polynucleotides of the present invention. Introduction of polynucleotides into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., 1986, *Basic Methods in Molecular Biology*, Elsevier, NY and Sambrook et al., 1989, *Molecular Cloning, 2nd Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY.

Production of the Subject Polypeptides

As hereinbefore mentioned, the present invention also teaches the production of polynucleotides of the invention by recombinant techniques.

The subject polynucleotides encode polypeptides which are the mature protein plus additional amino- or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. Generally, as is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

Thus, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

The polypeptides of the invention may be prepared by culturing the host/vector systems described above, in order to express the recombinant polypeptides. Recombinantly produced subject protein or parts thereof, may be further purified using techniques known in the art such as commercially available protein concentration systems, by salting out the protein followed by dialysis, by affinity chromatography, or using anion or cation exchange resins.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using DNA derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., *supra.*

Polynucleotides of the invention, encoding the heterologous structural sequence of a polynucleotide or polypeptide of the invention generally will be inserted into a vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates

translation of the polynucleotide or polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and the initiation codon. Also, generally, there will be a translation stop codon at the end of the expressed polynucleotide and there will be a polyadenylation signal in constructs for use in eukaryotic hosts. Transcription termination signal appropriately disposed at the 3' end of the transcribed region may also be included in the polynucleotide construct.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polynucleotide or polypeptide. These signals may be endogenous to the polynucleotide or they may be heterologous signals. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, or other such methods know to those skilled in the art. A subject polynucleotide or polypeptide can be recovered and purified from recombinant cell cultures by known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polynucleotide is denatured during isolation and or purification.

A nucleic acid molecule of the invention may be cloned into a glutathione S-transferase (GST) gene fusion system for example the pGEX-1T, pGEX-2T and pGEX-3X of Pharmacia. The fused gene may contain a strong *lac* promoter, inducible to a high level of expression by IPTG, as a regulatory element. Thrombin or factor Xa cleavage sites may be present which allow proteolytic cleavage of the desired polypeptide from the fusion product. The glutathione S-transferase-subject polypeptide fusion protein may be easily purified using a glutathione sepharose 4B column, for example from Pharmacia. The 26 kDa glutathione S-transferase polypeptide can be cleaved by thrombin (pGEX-1T or pGEX-2T) or factor Xa (pGEX-3X) and resolved from the polypeptide using the same affinity column. Additional chromatographic steps can be included if necessary, for example Sephadex or DEAE cellulose. The two enzymes may be monitored by protein and enzymatic assays and purity may be confirmed using SDS-PAGE.

The subject protein or parts thereof may also be prepared by chemical synthesis using techniques well known in the chemistry of proteins such as solid phase synthesis (Merrifield, 1964, *J. Am. Chem. Assoc.*, 85: 2149-2154) or synthesis in homogenous solution (Houbenweyl et al., 1987, *Methods of Organic Chemistry*, Wansch E. (ed), Vol. 15 I and II, Thieme, Germany).

Within the context of the present invention, the subject polypeptide includes various structural forms of the primary protein which retain biological activity. For example, the subject polypeptide may be in the form of acidic or basic salts or in neutral form. In addition, individual amino acid residues may be modified by oxidation or reduction. Furthermore, various substitutions, deletions or additions may be made to the amino acid or nucleic acid sequences, the net effect being that biological activity of the subject polypeptide is retained. Due to code degeneracy, for example, there may be considerable variation in nucleotide sequences encoding the same amino acid.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the carboxyl- or amino-terminus of the polypeptide to improve stability and persistence in the host cell during purification or during subsequent handling and storage. Also, fusion proteins may be added to the polynucleotide or polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polynucleotide or polypeptide. The addition of peptide moieties to polynucleotides or polypeptides to engender secretion or excretion, to improve stability or to facilitate purification, among others, are familiar and routine techniques in the art. In drug discovery, for example, proteins have been fused with antibody Fc portions for the purpose of high-throughput screening assays to identify antagonists (Bennett et al., 1995, *J. Mol. Recognit.*, 8: 52-58, and Johanson et al.,1995, *J. Biol. Chem.*, 270: 9459-9471).

Antibodies

With respect to protein-based testing, antibodies can be generated to the elongase gene product using standard immunological techniques, fusion proteins or synthetic peptides as described herein. Monoclonal antibodies can also be produced using now conventional techniques such as those described in Waldmann T.A., 1991, *Science*, 252: 1657-1662 and Harlow E. and Lane D. (eds.), 1988, *Antibodies: A Laboratory Manual*, Cold Harbour Press, Cold Harbour, NY. It

will also be appreciated that antibody fragments, i.e. Fab' fragments, can be similarly employed. Immunoassays, for example ELISAs, in which the test sample is contacted with antibody and binding to the gene product detected, can provide a quick and efficient method of determining the presence and quantity of the elongase gene product. For example, the antibodies can be used to test the effect of pharmaceuticals in subjects enrolled in clinical trials.

Thus, the present invention also provides polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, which are capable of specifically binding to the subject polypeptides and fragments thereof or to polynucleotide sequences from the subject polynucleotide region, particularly from the subject polypeptide locus or a portion thereof. The term "antibody" is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Polypeptides may be prepared synthetically in a peptide synthesizer and coupled to a carrier molecule (e.g., keyhole limpet hemocyanin) and injected over several months into rabbits. Rabbit sera is tested for immunoreactivity to the subject polypeptide or fragment. Monoclonal antibodies may be made by injecting mice with the protein polypeptides, fusion proteins or fragments thereof. Monoclonal antibodies are screened by ELISA and tested for specific immunoreactivity with subject polypeptide or fragments thereof (Harlow E. and Lane D. (eds.), 1988, *Antibodies: A Laboratory Manual*, Cold Harbour Press, Cold Harbour, NY). These antibodies are useful in assays as well as pharmaceuticals.

Once a sufficient quantity of desired polypeptide has been obtained, it may be used for various purposes. A typical use is the production of antibodies specific for binding. These antibodies may be either polyclonal or monoclonal, and may be produced by *in vitro* or *in vivo* techniques well known in the art. For production of polyclonal antibodies, an appropriate target immune system, typically mouse or rabbit, is selected. Substantially purified antigen is presented to the immune system in a fashion determined by methods appropriate for the animal and by other parameters well known to immunologists. Typical routes for injection are in footpads, intramuscularly, intraperitoneally, or intradermally. Of course, other species may be substituted for mouse or rabbit. Polyclonal antibodies are then purified using techniques known in the art, adjusted for the desired specificity.

An immunological response is usually assayed with an immunoassay. Normally, such immunoassays involve some purification of a source of antigen, for example, that produced by the same cells and in the same fashion as the antigen. A variety of immunoassay methods are well known in the art, such as in Harlow E. and Lane D. (eds.), 1988, *Antibodies: A*

Laboratory Manual, Cold Harbour Press, Cold Harbour, NY, or Goding J.W., 1996, *Monoclonal Antibodies: Principles and Practice: Production and Application of Monoclonal Antibodies in Cell Biology, Biochemistry and Immunology, 3rd edition*, Academic Press, NY.

Monoclonal antibodies with affinities of 10^8 M^{-1} or preferably 10^9 to 10^{10} M^{-1} or stronger will typically be made by standard procedures as described in Harlow E. and Lane D. (eds.), 1988, *Antibodies: A Laboratory Manual*, Cold Harbour Press, Cold Harbour, NY or Goding J.W., 1996, *Monoclonal Antibodies: Principles and Practice: Production and Application of Monoclonal Antibodies in Cell Biology, Biochemistry and Immunology, 3rd edition*, Academic Press, NY. Briefly, appropriate animals will be selected and the desired immunization protocol followed. After the appropriate period of time, the spleens of such animals are excised and individual spleen cells fused, typically, to immortalized myeloma cells under appropriate selection conditions. Thereafter, the cells are clonally separated and the supernatants of each clone tested for their production of an appropriate antibody specific for the desired region of the antigen.

Other suitable techniques involve *in vitro* exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors (Huse et al., 1989, *Science*, 246: 1275-1281). The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, magnetic particles and the like. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241. Also, recombinant immunoglobulins may be produced (see U.S. Pat. No. 4,816,567).

Generation of Polyclonal Antibody Against the Subject Polynucleotide

Sequences of the subject polynucleotide coding sequence are expressed as fusion protein in *E. coli*. The overexpressed protein is purified by gel elution and used to immunize rabbits and mice using a procedure similar to the one described by Harlow E. and Lane D. (eds.), 1988, *Antibodies: A Laboratory Manual*, Cold Harbour Press, Cold Harbour, NY. This procedure has been shown to generate antibodies against various other proteins (for example, see Kraemer et al., 1993, *J. Lipid Res.*, 34: 663-671).

Briefly, a stretch of coding sequence selected from the subject polynucleotide is cloned as a fusion protein in plasmid pET5A (Novagen, WI) or pMAL system (New England Biolabs, U.S.). After induction with IPTG, the overexpression of a fusion protein with the expected molecular weight is verified by SDS-PAGE. Fusion protein is purified from the gel by electroelution. The identification of the protein as the subject polypeptide fusion product can be verified by protein sequencing at the N-terminus. Next, the purified protein is used as immunogen in rabbits. Rabbits are immunized with 100 μg of the protein in complete Freund's adjuvant and boosted twice in 3 week intervals, first with 100 μg of immunogen in incomplete Freund's adjuvant followed by 100 μg of immunogen in PBS. Antibody containing serum is collected two weeks thereafter.

This procedure is repeated to generate antibodies against the mutant forms of the subject polypeptide. These antibodies, in conjunction with antibodies to wild type subject polypeptide, are used to detect the presence and the relative level of the mutant forms in various tissues and biological fluids.

Generation of Monoclonal Antibodies Specific for the Subject Polypeptide

Monoclonal antibodies are generated according to the following protocol. Mice are immunized with immunogen comprising intact subject polypeptide or its peptides (wild type or mutant) conjugated to keyhole limpet hemocyanin using glutaraldehyde or EDC as is well known.

The immunogen is mixed with an adjuvant. Each mouse receives four injections of 10 to 100 μg of immunogen and after the fourth injection blood samples are taken from the mice to determine if the serum contains antibody to the immunogen. Serum titer is determined by ELISA or RIA. Mice with sera indicating the presence of antibody to the immunogen are selected for hybridoma production.

Spleens are removed from immune mice and a single cell suspension is prepared as described by Harlow E. and Lane D. (eds.), 1988, *Antibodies: A Laboratory Manual*, Cold Harbour Press, Cold Harbour, NY. Cell fusions are performed essentially as described by Kohler G. and Milstein C., 1975, *Nature*, 256: 495-497. Briefly, P3.65.3 myeloma cells (American Type Culture Collection, Rockville, MD) are fused with immune spleen cells using polyethylene glycol as described by Harlow E. and Lane D. (eds.), 1988, *Antibodies: A Laboratory Manual*, Cold Harbour Press, Cold Harbour, NY. Cells are plated at a density of 2 x 10^5 cells/well in 96

well tissue culture plates. Individual wells are examined for growth and the supernatants of wells with growth are tested for the presence of subject polypeptide specific antibodies by ELISA or RIA using wild type or mutant target protein. Cells in positive wells are expanded and subcloned to establish and confirm monoclonality.

Clones with the desired specificities are expanded and grown as ascites in mice or in a hollow fiber system to produce sufficient quantities of antibody for characterization and assay development.

Sandwich Assay for the Subject Polypeptide

Monoclonal antibody is attached to a solid surface such as a plate, tube, bead, or particle. Preferably, the antibody is attached to the well surface of a 96-well ELISA plate. A 100 μl sample (e.g., serum, urine, tissue cytosol) containing the subject polypeptide/protein (wild-type or mutant) is added to the solid phase antibody. The sample is incubated for 2 hrs at room temperature. Next the sample fluid is decanted, and the solid phase is washed with buffer to remove unbound material. One hundred μl of a second monoclonal antibody (to a different determinant on the subject polypeptide/protein) is added to the solid phase. This antibody is labeled with a detector molecule or atom (e.g., ^{125}I, enzyme, fluorophore, or a chromophore) and the solid phase with the second antibody is incubated for two hrs at room temperature. The second antibody is decanted and the solid phase is washed with buffer to remove unbound material.

The amount of bound label, which is proportional to the amount of subject polypeptide/ protein present in the sample, is quantitated. Separate assays are performed using monoclonal antibodies which are specific for the wild-type subject polypeptide as well as monoclonal antibodies specific for each of the mutations identified in subject polypeptide.

Detecting Presence of or Predisposition for Disorders Affected by Lipid Metabolism and Monitoring Treatment of Same

As previously discussed, lipid metabolism is frequently disregulated in disease. It is likely that genetic polymorphisms in elongase genes will contribute to disease susceptibility.

The subject polynucleotides taught herein are useful to detect genetic polymorphisms of the

subject polynucleotides, or to detecting changes in the level of expression of the subject polynucleotides, as a diagnostic tool. Detection of an aberrant form of the subject polynucleotide, or a decrease or increase in the level of expression of the subject polynucleotide in a eukaryote, particularly a mammal, and especially a human, will provide a method for diagnosis of a disease. Eukaryotes (herein also "individual(s)"), particularly mammals, and especially humans, exhibiting genetic polymorphisms of the subject polynucleotides, or changes in expression of the subject polynucleotides may be detected by a variety of techniques.

Since elongase genes are widely expressed, test samples of the subject can be obtained from a variety of tissues including blood. An elongase gene test can also be included in panels of prenatal tests since elongase genes, DNA, RNA or protein can also be assessed in amniotic fluid. Quantitative testing for elongase gene transcript and gene product is thus also contemplated within the scope of the present invention.

Nucleic acid and protein-based methods for screening genetic polymorphisms in elongase genes are all within the scope of the present teachings. For example, knowing the sequence of the elongase gene, DNA or RNA probes can be constructed and used to detect mutations in elongase genes through hybridization with genomic DNA in a tissue such as blood using conventional techniques. RNA or cDNA probes can be similarly probed to screen for mutations in elongase genes or for quantitative changes in expression. A mixture of different probes, i.e. "probe cocktail", can also be employed to test for more than one mutation.

With respect to nucleic acid-based testing, genomic DNA may be used directly for detection of a specific sequence or may be amplified enzymatically *in vitro* by using PCR prior to analysis (Saiki et al., 1985, *Science*, 230: 1350-1353 and Saiki et al., 1986, *Nature*, 324: 163-166). Reviews of this subject have been presented by Caskey C.T., 1989, *Science*, 236: 1223-1229 and by Landegren et al., 1989, *Science*, 242: 229-237. The detection of specific DNA sequence may be achieved by methods such as hybridization using specific oligonucleotides (Wallace et al., 1986, *Cold Spring Harbour Symp. Quant. Biol.*, 51: 257-261), direct DNA sequencing (Church et al., 1988, *Proc. Natl. Acad. Sci.*, 81: 1991-1995, the use of restriction enzymes (Flavell et al., 1978, *Cell*, 15: 25-41; Geever et al., 1981, *Proc. Natl. Acad. Sci.*, 78: 5081-5085), discrimination on the basis of electrophoretic mobility in gels with denaturing reagent (Myers et al., 1986, *Cold Spring Harbour Sym. Quant. Biol.*, 51: 275-284), RNase protection (Myers et al., 1985, *Science*, 230: 1242-1246), chemical cleavage (Cotton et al., 1985, *Proc. Natl. Acad. Sci.*, 85: 4397-4401), and the ligase-mediated detection procedure (Landegren et al., 1988, *Science*, 241: 1077-1080). Using PCR, characterization of the level of

or condition of the subject polynucleotides present in the individual may be made by comparative analysis.

With respect to protein-based testing, antibodies can be generated to the elongase gene product using standard immunological techniques, fusion proteins or synthetic peptides as described herein.

With the characterization of the elongase gene product and its function, functional assays can also be used for elongase gene diagnosis and screening and to monitor treatment. For example, enzymatic testing to determine levels of gene function, rather than direct screening of the elongase gene or product, can be employed. Testing of this nature has been utilized in other diseases and conditions, such as in Tay-Sachs.

The invention thus provides a process for detecting disease by using methods known in the art and methods described herein to detect changes in expression of or mutations to the subject polynucleotides. For example, decreased expression of a subject polynucleotide can be measured using any one of the methods well known in the art for the quantification of polynucleotides, such as, for example, PCR, RT-PCR, DNase protection, Northern blotting and other hybridization methods. Thus, the present invention provides a method for detecting disorders affected by lipid metabolism, and a method for detecting a genetic pre-disposition for such diseases including eczema, cardiovascular disorders (including but not limited to hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease, cerebrovascular disease and peripheral vascular disease), inflammation (including but not limited to sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne), body weight disorders (including but not limited to obesity, cachexia and anorexia), psychiatric disorders, cancer, cystic fibrosis, pre-menstrual syndrome, diabetes and diabetic complications.

Drug Screening Assays

The present teachings provide methods for screening compounds to identify those which enhance (agonist) or block (antagonist) the action of subject polypeptides or polynucleotides, such as its interaction with fatty acid binding molecules. The identification of the subject polynucleotides in inherited fatty acid disorders, combined with advances in the field of transgenic methods, provides the information necessary to further study human diseases. This is extraordinarily useful in modeling familial forms of fatty acid disorders and other diseases of fatty acid metabolism including eczema, cardiovascular disorders (including but not limited to hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease,

cerebrovascular disease and peripheral vascular disease), inflammation (including but not limited to sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne), body weight disorders (including but not limited to obesity, cachexia and anorexia), psychiatric disorders, cancer, cystic fibrosis, pre-menstrual syndrome, diabetes and diabetic complications. Drug screening assays are made effective by use of the control regions of the genes described in the present invention or part of it, in a yeast based DNA-protein interaction assay (yeast one-hybrid). The use of the genes described here, or parts thereof, or the transcribed RNA in a yeast protein-protein interaction (2-hybrid) or protein-RNA interaction assays for drug screening also provide effective drug screening methods. Such interacting molecules can also be reconstructed *in vitro* for drug screening purposes.

For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, may be prepared from a cell that expresses a molecule that binds a subject polynucleotide. The preparation is incubated with labeled polynucleotide in the absence or the presence of a candidate molecule which may be an agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand.

Effects of potential agonists and antagonists may by measured, for instance, by determining activity of a reporter system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect to a baseline (control) measurement. Reporter systems that may be useful in this regard include, but are not limited to, colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in elongase enzyme activity, and binding assays known in the art.

Another example of an assay for antagonists is a competitive assay that combines a subject polypeptide and a potential antagonist with membrane-bound subject polypeptide-binding molecules, recombinant subject polypeptide binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. A subject polypeptide can be labeled, such as by radioactivity or a colorimetric compound, such that the number of subject polypeptide molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide or polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, peptides, polypeptides,

such as closely related proteins or antibodies that bind the same sites on a binding molecule, without inducing subject polypeptide-induced activities, thereby preventing the action of the subject polypeptide by excluding the subject polypeptide from binding. Potential antagonists include antisense molecules (Okano et al., 1988, *EMBO J.*, 7: 3407-3412). Potential antagonists include compounds related to and derivatives of the subject polypeptides.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide or polypeptide of the invention and thereby inhibit or extinguish its activity. Potential agonists may be selected from the group consisting of small organic molecules, peptides, polypeptides, antisense molecules, oligonucleotides, polynucleotides, fatty acids, and chemical and functional derivatives thereof.

Developing modulators of the biological activities of specific elongases requires differentiating elongase isozymes present in a particular assay preparation. The classical enzymological approach of isolating elongases from natural tissue sources and studying each new isozyme may be used. Another approach has been to identify assay conditions which might favor the contribution of one isozyme and minimize the contribution of others in a preparation. Still another approach is the separation of elongases by immunological means. Each of the foregoing approaches for differentiating elongase isozymes is time consuming. As a result many attempts to develop selective elongase modulators have been performed with preparations containing more than one isozyme. Moreover, elongase preparations from natural tissue sources are susceptible to limited proteolysis and may contain mixtures of active proteolytic products that have different kinetic, regulatory and physiological properties than the full length elongases.

Recombinant subject polypeptide products of the invention greatly facilitate the development of new and specific modulators. The need for purification of an isozyme can be avoided by expressing it recombinantly in a host cell that lacks endogenous elongase activity. Once a compound that modulates the activity of the elongase is discovered, its selectivity can be evaluated by comparing its activity on the particular subject enzyme to its activity on other elongase isozymes. Thus, the combination of the recombinant subject polypeptide products of the invention with other recombinant elongase products in a series of independent assays provides a system for developing selective modulators of particular elongases. Selective modulators may include, for example, antibodies and other proteins or peptides which specifically bind to the subject polypeptide or polynucleotide, oligonucleotides which specifically bind to the subject polypeptide (see Patent Cooperation Treaty International Publication No. WO 93/05182 which describes methods for selecting oligonucleotides which

selectively bind to target biomolecules) or the subject polynucleotide (e.g., antisense oligonucleotides) and other non-peptide natural or synthetic compounds which specifically bind to the subject polynucleotide or polypeptide. Mutant forms of the subject polynucleotide which alter the enzymatic activity of the subject polypeptide or its localization in a cell are also contemplated. Crystallization of recombinant subject polypeptides alone and bound to a modulator, analysis of atomic structure by X-ray crystallography, and computer modeling of those structures are methods useful for designing and optimizing non-peptide selective modulators. See, for example, Erickson et al., 1992, *Ann. Rep. Med. Chem.*, 27: 271-289 for a general review of structure-based drug design.

Targets for the development of selective modulators include, for example: (1) the regions of the subject elongases which contact other proteins and/or localize the proteins within a cell, (2) the regions of the proteins which bind substrate, and (3) the phosphorylation site(s) of the subject polypeptides.

Thus, the present invention provides methods for screening and selecting compounds which promote disorders affected by lipids. As well, the present invention provides methods for screening and selecting compounds which treat or inhibit progression of diseases associated with lipid metabolism, such as eczema, cardiovascular disorders (including but not limited to hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease, cerebrovascular disease and peripheral vascular disease), inflammation (including but not limited to sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne), body weight disorders (including but not limited to obesity, cachexia and anorexia), psychiatric disorders, cancer, cystic fibrosis, pre-menstrual syndrome, diabetes and diabetic complications, and other diseases not necessary related to lipid metabolism.

Protein Interaction Assays for DNA control regions, CDS and RNA of Elongase Genes.

Protein interaction is implicated in virtually every biological process in the cell, for example, metabolism, transport, signaling and disease. Development of the yeast 2-hybrid and 1-hybrid systems have made it possible to study and identify protein-protein interaction, protein-DNA interaction or protein-RNA interaction *in vivo* (Fields S. and Song O., 1989, *Nature*, 340: 245-246; Ulmasov et al., 1997, *Science*, 276: 1865-1868; Furuyama K. and Sassa S., 2000, *J. Clin. Invest.*, 105: 757-764 and Gyuris et al., 1993, *Cell*, 75: 791-803). Because these interactions are key to cellular functions, identification of interacting partners is the first step towards elucidation of function and involvement in pathogenesis. New chemical entities that modulate (inhibit or activate) such interactions may have strong pharmaceutical and therapeutic benefit

in human, animal as well as plant diseases. It is now known that in sideroblastic anemic patients, the interaction between succinyl-CoA synthetase and the heme biosynthetic enzyme δ-aminolevulinate synthase-E (ALAS-E) is disrupted (Furuyama K. and Sassa S., 2000, *J. Clin. Invest.*, 105: 757-764). Inhibition of gene expression in human cells through small molecule-RNA interaction have been recently described (Hwang et al., 1999, *Proc. Natl. Acad. Sci.*, 96: 12997-13002). The use of protein-RNA inhibition technology is a potential approach for development of anti-HIV therapeutics (Hamy et al., 1997, *Proc. Natl. Acad. Sci.*, 94: 3548-3553 and Mei et al., 1998, *Biochemistry*, 37: 14204-14212).

Drug Design

Antagonists and agonists and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds. The pharmaceutical compositions may be administered in any effective, convenient manner. including, for instance, administration by direct microinjection into the affected area, or by intravenous or other routes. These compositions of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a medium additive or a therapeutically effective amount of antagonists or agonists of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation is prepared to suit the mode of administration.

Modulation of elongase gene function can be accomplished by the use of therapeutic agents or drugs which can be designed to interact with different aspects of elongase structure or function. For example, a drug or antibody can bind to a structural fold of the protein to correct a defective structure. Alternatively, a drug might bind to a specific functional residue and increase its affinity for a substrate or cofactor. Efficacy of a drug or agent can be identified by a screening program in which modulation is monitored *in vitro* in cell systems in which a defective elongase is expressed.

Alternatively, drugs can be designed to modulate the activity of proteins of elongase genes from knowledge of the structure and function correlations for such proteins and from knowledge of the specific defect in various mutant proteins (Copsey et al., 1988, *Genetically Engineered Human Therapeutic Drugs*, Stockton Press, NY).

Gene Therapy

A variety of gene therapy approaches may be used in accordance with the invention to modulate expression of the subject polynucleotides *in vivo*. For example, antisense DNA molecules may be engineered and used to block translation of mRNA of the subject polynucleotides *in vivo*. Alternatively, ribozyme molecules may be designed to cleave and destroy the mRNA of the subject polynucleotides *in vivo*. In another alternative, oligonucleotides designed to hybridize to the 5' region of the subject polynucleotide (including the region upstream of the coding sequence) and form triple helix structures may be used to block or reduce transcription of the subject polynucleotide. In yet another alternative, nucleic acid encoding the full length wild-type subject polynucleotide may be introduced *in vivo* into cells which otherwise would be unable to produce the wild-type subject polynucleotide product in sufficient quantities or at all.

For example, in conventional replacement therapy, gene product or its functional equivalent is provided to the patient in therapeutically effective amounts. Elongases can be purified using conventional techniques such as those described in Deutcher M. (ed.), 1990, *Guide to Protein Purification, Meth. Enzymol.*, Vol. 182. Sufficient amounts of gene product or protein for treatment can be obtained, for example, through cultured cell systems or synthetic manufacture. Drug therapies which stimulate or replace the gene product can also be employed. Delivery vehicles and schemes can be specifically tailored to the particular target gene.

Gene therapy using recombinant technology to deliver the gene into the patient's cells, or vectors which will supply the patient with gene product *in vivo*, is also within the scope of the invention. Retroviruses have been considered preferred vectors for experiments in somatic gene therapy, with a high efficiency of infection and stable integration and expression (Orkin et al., 1988, *Prog. Med. Genet.*, 7: 130-142). For example, elongase cDNAs can be cloned into a retroviral vector and driven from either its endogenous promoter or from the retroviral LTR (long terminal repeat). Other delivery systems which can be utilized include adeno-associated virus (McLaughlin et al., 1988, *J. Virol.*, 62: 1963-1973), vaccinia virus (Moss et al., 1987, *Annu. Rev. Immunol.*, 5: 305-324), bovine papilloma virus (Rasmussen et al., 1987, *Meth. Enzymol.*, 139: 642-654), or a member of the herpes virus group such as Epstein-Barr virus (Margolskee et al., 1988, *Mol. Cell. Biol.*, 8: 2837-2847).

Antisense, ribozyme and triple helix nucleotides are designed to inhibit the translation or transcription of the subject polynucleotides. To accomplish this, the oligonucleotides used should be designed on the basis of relevant sequences unique to the subject polynucleotides. For example, and not by way of limitation, the oligonucleotides should not fall within those regions where the nucleotide sequence of a subject polynucleotide is most homologous to that of other polynucleotides, herein referred to as "unique regions".

In the case of antisense molecules, it is preferred that the sequence be chosen from the unique regions. It is also preferred that the sequence be at least 18 nucleotides in length in order to achieve sufficiently strong annealing to the target mRNA sequence to prevent translation of the sequence (Izant J.G. and Weintraub H., 1984, *Cell*, 36: 1007-1015 and Rosenberg et al., 1985, *Nature*, 313: 703-706).

In the case of the "hammerhead" type of ribozymes, it is also preferred that the target sequences of the ribozymes be chosen from the unique regions. Ribozymes are RNA molecules which possess highly specific endoribonuclease activity. Hammerhead ribozymes comprise a hybridizing region which is complementary in nucleotide sequence to at least part of the target RNA, and a catalytic region which is adapted to cleave the target RNA. The hybridizing region contains 9 or more nucleotides. Therefore, the hammerhead ribozymes of have a hybridizing region which is complementary to the sequences listed above and is at least nine nucleotides in length. The construction and production of such ribozymes are well known in the art and are described more fully in Haseloff J. and Gerlach W.L., 1988, *Nature*, 334: 585-591.

The ribozymes also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug et al., 1984, *Science*, 224: 574-578; Zaug A.J. and Cech T.R, 1986, *Science*, 231: 470-475; Zaug et al., 1986, *Nature*, 324: 429-433; Patent Publication Treaty International Patent Application No. WO 88/04300 and Been M.D. and Cech T.R., 1986, *Cell*, 47: 207-216). The Cech endoribonucleases have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. Cech-type ribozymes target eight base-pair active site sequences are present in a subject polynucleotide but not other polynucleotides for elongases.

The compounds can be administered by a variety of methods which are known in the art, including, but not limited to the use of liposomes as a delivery vehicle. Naked DNA or RNA molecules may also be used where they are in a form which is resistant to degradation, such as by modification of the ends, by the formation of circular molecules, or by the use of alternate bonds including phosphothionate and thiophosphoryl modified bonds. In addition, the delivery of nucleic acid may be by facilitated transport where the nucleic acid molecules are conjugated to polylysine or transferrin. Nucleic acid may also be transported into cells by any of the various viral carriers, including but not limited to, retrovirus, vaccinia, adeno-associated virus, and adenovirus.

Alternatively, a recombinant nucleic acid molecule which encodes, or is, such antisense, ribozyme, triple helix, or subject polynucleotide molecule can be constructed. This nucleic acid molecule may be either RNA or DNA. If the nucleic acid encodes an RNA, it is preferred that the sequence be operatively attached to a regulatory element so that sufficient copies of the desired RNA product are produced. The regulatory element may permit either constitutive or regulated transcription of the sequence. A transfer vector such as a bacterial plasmid or viral RNA or DNA, encoding one or more of the RNAs, may be transfected into cells or cells of an organism (Llewellyn et al., 1987, *J. Mol. Biol.*, 195: 115-123 and Hanahan et al., 1983, *J. Mol. Biol.*, 166: 557-580). Once inside the cell, the transfer vector may replicate, and be transcribed by cellular polymerases to produce the RNA or it may be integrated into the genome of the host cell. Alternatively, a transfer vector containing sequences encoding one or more of the RNAs may be transfected into cells or introduced into cells by way of micromanipulation techniques such as microinjection, such that the transfer vector or a part thereof becomes integrated into the genome of the host cell.

Composition, Formulation, and Administration of Pharmaceutical Compositions

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained by solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol, or cellulose preparations such as, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone. If desired, disintegrating agents may be added, such as the cross-linked polyvinylpyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration.

For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges (e.g. gelatin) for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multidose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for

example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied.

Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing therapeutic agent. Various sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include, but are not limited to, calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

Many of the compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Pharmaceutically compatible salts may be formed with many acids, including but, not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms.

Suitable routes of administration may, for example, include oral, rectal, transmucosal, transdermal, or intestinal administration; or parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into an affected area, often in a depot or sustained release formulation.

Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with an antibody specific for affected cells. The liposomes will be targeted to and taken up selectively by the cells.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. It is appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like. In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example, as a sterile aqueous dispersion, preferably isotonic. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms associated with such disorders. Techniques for formulation and administration of the compounds of the instant application may be found in "*Remington's Pharmaceutical Sciences*," Mack Publishing Co., Easton, Pa., latest edition. For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.001 mg/kg to 10 mg/kg, typically around 0.01 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The compounds of the invention may be particularly useful in animal disorders (veterinarian indications), and particularly mammals.

The invention further provides diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

DEFINITIONS

To facilitate a complete understanding of the invention, the terms defined below have the following meaning:

Agonist refers to any molecule or pharmaceutical agent, such as a drug or hormone, which enhances the activity of another molecule.

Antagonist refers to any molecule or pharmaceutical agent, such as a drug or hormone, which inhibits or extinguishes the activity of another molecule.

Chemical Derivative. As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties can improve the molecule's solubility, absorption, biological half life, and the like. The moieties can alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, and the like. Moieties capable of mediating such effects are disclosed in Mack E.W., 1990, *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Easton, Pa., 13th edition. Procedures for coupling such moieties to a molecule are well known in the art.

Compositions include genes, proteins, polynucleotides, peptides, compounds, drugs, and pharmacological agents.

Control region refers to a nucleic acid sequence capable of, or required for, assisting or impeding initiation, termination, or otherwise regulating the transcription of a gene. The control region may include a promoter, enhancer, silencer and/or any other regulatory element. A control region also includes a nucleic acid sequence that may or may not be independently or exclusively sufficient to initiate, terminate, or otherwise regulate transcription, however, is capable of effecting such regulation in association with other nucleic acid sequences.

Desaturase refers to a fatty acid desaturase, which is an enzyme capable of generating a double bond in the hydrocarbon region of a fatty acid molecule.

Disorder as used herein refers to derangement or abnormality of structure or function. Disorder includes disease.

Drug. Drugs include, but are not limited to proteins, peptides, degenerate peptides, agents purified from conditioned cell medium, organic molecules, inorganic molecules, antibodies or oligonucleotides. The drug can be naturally occurring or synthetically or recombinantly produced.

Enhancer is a nucleic acid sequence comprising a DNA regulatory element that enhances or increases transcription when bound by a specific transcription factor or factors. Moreover, an enhancer may function in either orientation and in any location (upstream or downstream relative to the promoter) to effect and generate increased levels of gene expression when bound by specific factors. In addition, according to the present invention, an enhancer also refers to a compound (i.e. test compound) that increases or promotes the enzymatic activity of the elongase gene, and/or increases or promotes the transcription of the gene.

Fatty Acids are a class of compounds comprising a long saturated or mono or polyunsaturated hydrocarbon chain and a terminal carboxyl group.

Fatty Acid Delta-5-Desaturase (D5D) is an enzyme capable of generating a double bond between carbons 5 and 6 from the carboxyl group in a fatty acid molecule.

Fatty Acid Delta-6-Desaturase is an enzyme capable of generating a double bond between carbons 6 and 7 from the carboxyl group in a fatty acid molecule.

Fatty Acid Elongase is an enzyme required for the addition of an acetyl group or a 2-carbon chain to the carboxyl end of a fatty acid.

Functional Enzyme, as used herein, refers to a biologically active or non-active protein with a known enzymatic activity.

Functional Derivative. A "functional derivative" of a sequence, either protein or nucleic acid, is a molecule that possesses a biological activity (either functional or structural) that is substantially similar to a biological activity of the protein or nucleic acid sequence. A functional derivative of a protein can contain post-translational modifications such as covalently linked carbohydrate, depending on the necessity of such modifications for the performance of a specific function. The term "functional derivative" is intended to include the "fragments," "sequences," "variants," "analogs," or "chemical derivatives" of a molecule.

Gene refers to a nucleic acid molecule or a portion thereof, the sequence of which includes information required for the production of a particular protein or polypeptide chain. The polypeptide can be encoded by a full-length sequence or any portion of the coding sequence, so long as the functional activity of the protein is retained. A gene may comprise regions preceding and following the coding region as well as intervening sequences (introns) between individual coding sequences (exons). A **"heterologous"** region of a nucleic acid construct (i.e. a heterologous gene) is an identifiable segment of DNA within a larger nucleic acid construct that is not found in association with the other genetic components of the construct in nature. Thus, when the heterologous gene encodes a mammalian elongase gene, the gene will usually be flanked by a promoter that does not flank the structural genomic DNA in the genome of the source organism.

Host system may comprise a cell, tissue, organ, organism or any part thereof, which provides an environment or conditions that allow for, or enable, transcription and/or transcription.

Identity, similarity, homology or homologous, refer to relationships between two or more polynucleotide sequences, as determined by comparing the sequences. In the art, identity also means the degree of sequence relatedness between polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. Both identity and similarity can be readily calculated (Lesk A.M., ed., 1988, *Computational Molecular Biology*, Oxford University Press, NY; Smith D.W., ed., 1993, *Biocomputing: Informatics and Genome Project*, Academic Press, NY; Griffin A.M. and Griffin H.G., eds., 1994, *Computer Analysis of Sequence Data, Part I*, Humana Press, NJ; von Heijne G., 1987, *Sequence Analysis in Molecular Biology*, Academic Press, NY and Gribskov M. and Devereux J., eds., 1991, *Sequence Analysis Primer*, M Stockton Press, NY). While there exist a number of methods to measure identity and similarity between two polynucleotide sequences, both terms are well known to skilled artisans (von Heijne G., 1987, *Sequence Analysis in Molecular Biology*, Academic Press, NY; Gribskov M. and Devereux J., eds., 1991, *Sequence Analysis Primer*, M Stockton Press, NY and Carillo H. and Lipman D., 1988, *SIAM J. Applied Math.*, 48: 1073). Methods commonly employed to determine identity or similarity between sequences include, but are not limited to those disclosed in Carillo H. and Lipman D., 1988, *SIAM J. Applied Math.*, 48: 1073. Methods to determine identity and similarity are codified in computer programs. Computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux et al., 1984, *Nucl. Acid Res.*, 12: 387-395), BLASTP, BLASTN and FASTA (Altschul et al., 1990, *J. Molec. Biol.*, 215: 403-410).

Isolated means altered "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a naturally occurring polynucleotide naturally present in a living organism in its natural state is not "isolated," but the same polynucleotide separated from coexisting materials of its natural state is "isolated", as the term is employed herein. As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNA, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in culture or in whole organisms, such DNA still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides may occur in a composition, such as a media formulations, solutions for introduction of polynucleotides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides within the meaning of that term as it is employed herein.

Mutation. A "mutation" is any detectable change in the genetic material. A mutation can be any (or a combination of) detectable, unnatural change affecting the chemical or physical constitution, mutability, replication, phenotypic function, or recombination of one or more deoxyribonucleotides; nucleotides can be added, deleted, substituted for, inverted, or transposed to new positions with and without inversion. Mutations can occur spontaneously and can be induced experimentally by application of mutagens or by site-directed mutagenesis. A mutant polypeptide can result from a mutant nucleic acid molecule.

Nucleic acid construct refers to any genetic element, including, but not limited to, plasmids and vectors, that incorporate polynucleotide sequences. For example, a nucleic acid construct may be a vector comprising a promoter or control region that is operably linked to a heterologous gene.

Operably linked as used herein indicates the association of a promoter or control region of a nucleic acid construct with a heterologous gene such that the presence or modulation of the promoter or control region influences the transcription of the heterologous gene, including genes for reporter sequences. Operably linked sequences may also include two segments that are transcribed onto the same RNA transcript. Thus, two sequences, such as a promoter and a "reporter sequence" are operably linked if transcription commencing in the promoter produces an RNA transcript of the reporter sequence.

Plasmids. Starting plasmids disclosed herein are either commercially available, publicly available, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention.

Polynucleotides(s) of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be double-stranded or single-stranded. Single-stranded polynucleotides may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand. Polynucleotides generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded, or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term polynucleotide also includes DNA or DNA that contain one or more modified bases. Thus, DNA or DNA with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNA or DNA comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, *inter alia*. Polynucleotides embraces short polynucleotides often referred to as oligonucleotide(s). It will also be appreciated that RNA made by transcription of this doubled stranded nucleotide sequence, and an antisense strand of a nucleic acid molecule of the

invention or an oligonucleotide fragment of the nucleic acid molecule, are contemplated within the scope of the invention. An antisense sequence is constructed by inverting the sequence of a nucleic acid molecule of the invention, relative to its normal presentation for transcription. Preferably, an antisense sequence is constructed by inverting a region preceding the initiation codon or an unconserved region. The antisense sequences may be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art.

Promoter refers to a nucleic acid sequence comprising a DNA regulatory element capable of binding RNA polymerase directly or indirectly to initiate transcription of a downstream (3' direction) gene. In accordance with the present invention, a promoter of a nucleic acid construct that includes a nucleotide sequence, wherein the nucleotide sequence may be linked to a heterologous gene such that the induction of the promoter influences the transcription of the heterologous gene.

Purified: A "purified" protein or nucleic acid is a protein or nucleic acid preparation that is generally free of contaminants, whether produced recombinantly, chemically synthesized or purified from a natural source.

Recombinant refers to recombined or new combinations of nucleic acid sequences, genes, or fragments thereof which are produced by recombinant DNA techniques and are distinct from a naturally occurring nucleic acid sequence

Regulatory element refers to a deoxyribonucleotide sequence comprising the whole, or a portion of, a nucleic acid sequence to which an activated transcriptional regulatory protein, or a complex comprising one or more activated transcriptional regulatory proteins, binds so as to transcriptionally modulate the expression of an associated gene or genes, including heterologous genes.

Reporter gene is a nucleic acid coding sequence whose product is a polypeptide or protein that, is not otherwise produced by the host cell or host system, or which is produced in minimal or negligible amounts in the host cell or host system, and which is detectable by various known methods such that the reporter gene product may be quantitatively assayed to analyse the level of transcriptional activity in a host cell or host system. Examples include genes for luciferase, chloramphenicol acetyl transferase (CAT), beta-galactosidase, secreted placental alkaline phosphatase and other secreted enzymes.

Silencer refers to a nucleic acid sequence or segment of a DNA control region such that the presence of the silencer sequence in the region of a target gene suppresses the transcription of the target gene at the promoter through its actions as a discrete DNA segment or through the actions of trans-acting factors that bind to these genetic elements and consequently effect a negative control on the expression of a target gene.

Stringent hybridization conditions are those which are stringent enough to provide specificity, reduce the number of mismatches and yet are sufficiently flexible to allow formation of stable hybrids at an acceptable rate. Such conditions are known to those skilled in the art and are described, for example, in Sambrook et al., 1989, *Molecular Cloning, 2nd Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY or Ausubel et al., 1994-, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY. By way of example only, stringent hybridization with short nucleotides may be carried out at 5-10°C below the T_M using high concentrations of probe such as 0.01-1.0 pmole/ml. Preferably, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

Tag refers to a specific short amino acid sequence, or the oligonucleotide sequence that encodes it, wherein said amino acid or nucleic acid sequence may comprise or encode, for example, a c-myc epitope and/or a string of six histidine residues recognizable by commercially available antibodies. In practice, a tag facilitates the subsequent identification and purification of a tagged protein.

Tagged protein as used herein refers to a protein comprising a linked tag sequence. For example, a tagged protein includes a mammalian elongase polypeptide linked to a c-myc epitope and six histidine residues at the carboxyl terminus of the amino acid sequence.

Test compounds as used herein encompass small molecules (e.g. small organic molecules), pharmacological compounds or agents, peptides, proteins, antibodies or antibody fragments, and nucleic acid sequences, including DNA and RNA sequences.

Transfection refers to a process whereby exogenous or heterologous DNA (i.e. a nucleic acid construct) is introduced into a recipient eukaryotic host cell. Therefore, in eukaryotic cells, the acquisition of exogenous DNA into a host cell is referred to as transfection. In prokaryotes and eukaryotes (for example, yeast and mammalian cells) introduced DNA may be maintained on an episomal element such as a plasmid or integrated into the host genome. With respect to eukaryotic cells, a stably transfected cell is one in which the introduced DNA has become

integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the introduced DNA.

Transformation refers to a process whereby exogenous or heterologous DNA (i.e. a nucleic acid construct) is introduced into a recipient prokaryotic host cell. Therefore, in prokaryotic cells, the acquisition of exogenous DNA into a host cell is referred to as transformation. Transformation in eukaryotes refers to the conversion or transformation of eukaryotic cells to a state of unrestrained growth in culture, resembling a tumorigenic condition. In prokaryotes and eukaryotes (for example, yeast and mammalian cells) introduced DNA may be maintained on an episomal element such as a plasmid or integrated into the host genome. With prokaryotic cells, a stably transformed bacterial cell is one in which the introduced DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the prokaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the introduced DNA.

Transfection/transformation as used herein refers to a process whereby exogenous or heterologous DNA (e.g. a nucleic acid construct) has been introduced into a eukaryotic or prokaryotic host cell or into a host system.

Variant(s) of polynucleotides are polynucleotides that differ in nucleotide sequence from another, reference polynucleotide. A "variant" of a protein or nucleic acid is meant to refer to a molecule substantially similar in structure and biological activity to either the protein or nucleic acid. Thus, provided that two molecules possess a common activity and can substitute for each other, they are considered variants as that term is used herein even if the composition or secondary, tertiary, or quaternary structure of one of the molecules is not identical to that found in the other, or if the amino acid or nucleotide sequence is not identical. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical. Changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide or polynucleotide with the same amino acid sequence as the reference. Changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide or polynucleotide encoded by the reference sequence.

Vector. A plasmid or phage DNA or other DNA sequence into which DNA can be inserted to be cloned. The vector can replicate autonomously in a host cell, and can be further characterized by one or a small number of endonuclease recognition sites at which such DNA sequences can be cut in a determinable fashion and into which DNA can be inserted. The vector can further contain a marker suitable for use in the identification of cells transformed with the vector. Markers, for example, are tetracycline resistance or ampicillin resistance. The words "cloning vehicle" are sometimes used for "vector."

The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims. Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

The present invention is further described and will be better understood by referring to the working examples set forth below. These non-limiting examples are to be considered illustrative only of the principles of the invention. Since numerous modifications and changes will readily occur to those skilled in the art, it is not desired to limit the invention to the exact construction and operation shown and described. Accordingly, all suitable modifications and equivalents may be used and will fall within the scope of the invention and the appended claims.

EXAMPLES

The present invention is further described by the following examples. These examples, while illustrating certain specific aspects of the invention, do not portray the limitations or circumscribe the scope of the disclosed invention.

Example 1- Cloning ELG1

ELG1 was cloned into the pYES2/CT yeast expression vector (Invitrogen) using PCR. Two plasmid constructions were made for the production of the ELG1 protein with either a C-terminal tag containing the V-5 epitope and polyhistidine peptide (ELG1/V5-His), or the ELG1 protein without the tag (ELG1). The forward primer (5'-CACGCGGGTACCAGGATGGAGGCTGTTGTGAAC-3') contains the translation start codon and a *Kpn*I site (underlined). The reverse primers for cloning ELG1 and ELG1/V5-His,

5'-ATATCACGATGCGGCCGCTCAGTTGGCCTTGACCTTGGC-3' and 5'-ATATCACGATGCGGCCGCCAGTTGGCCTTGACCTTGGC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG1 provides the translation stop codon. The reverse primer for cloning ELG1/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Advantage-HF polymerase (Clontech) as per the manufacturer's instructions. The SuperScript human leukocyte cDNA library (Gibco BRL) was used as the DNA template for cloning ELG1. pTh1009.1 (defined below) was used as the template for cloning ELG1/V5-His.

The PCR products were gel purified, digested with *Kpn*I and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transform *E. coli* strain INVαF' (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG1 and ELG1/V5-His were designated pTh1009.1 (Figure 7) and pTh1009.2 (Figure 18), respectively.

Example 2 - Cloning ELG2

Obtaining Complete Coding Sequence for ELG2

Clones containing the complete coding sequence for ELG2 were obtained from the SuperScript human leukocyte cDNA library (Gibco BRL) using the GeneTrapper cDNA Positive Selection System (Gibco BRL) as per the manufacturer's instructions. The sequence of the oligonucleotide used to probe the library and repair the captured cDNA target was 5'-GTAACAGGAGTATGGGAAGGCA-3'. The repaired DNA was used to transform UltraMax DH5α-FT cells (Gibco BRL). Clones containing ELG2 were identified by colony PCR using 5'-TTGGACTCACACTGCTGTCTCT-3' and 5'-GTGTGGCACCAAAATAAGAGTG-3' as gene specific primers and Platinum Taq DNA polymerase (Gibco BRL). Plasmid DNA was isolated from selected colonies and their inserts were sequenced. The nucleotide sequence obtained was used to identify the open reading frame for ELG2 and to design primers for cloning ELG2 into a yeast expression vector. A plasmid containing the complete ELG2 coding sequence was designated pSh1010.1.

Cloning ELG2 into Expression Vector

ELG2 was cloned into the pYES2/CT yeast expression vector (Invitrogen) using PCR. Two plasmid constructions were made for the production of the ELG2 protein with either a C-

terminal tag containing the V-5 epitope and polyhistidine peptide (ELG2/V5-His), or the ELG2 protein without the tag (ELG2). The forward primer (5'-CACGCGGGATCCCAAATGGAACATTTTGATGCATCAC-3') contains the translation start codon and a *Bam*HI site (underlined). The reverse primers for cloning ELG2 and ELG2/V5-His, 5'-ATATCACGATGCGGCCGCTCAATCCTTCCGCAGCTTCC-3' and 5'-ATATCACGATGCGGCCGCCAATCCTTCCGCAGCTTCC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG2 provides the translation stop codon. The reverse primer for cloning ELG2/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Advantage-HF polymerase (Clontech) as per the manufacturer's instructions. pSh1010.1 was used as the DNA template for cloning ELG2. pMr1014.1 (described below) was used as the DNA template for ELG2/V5-His.

The PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transform *E. coli* strain TOP10F' (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG2 and ELG2/V5-His were designated pTh1014.1 and pTh1014.2, respectively.

Example 3- Cloning ELG3

ELG3 was cloned into the pYES2/CT yeast expression vector (Invitrogen) using PCR. Two plasmid constructions were made for the production of the ELG3 protein with either a C-terminal tag containing the V-5 epitope and polyhistidine peptide (ELG3/V5-His), or the ELG3 protein without the tag (ELG3). The forward primer (5'-CACGCGGGATCCATCATGGAACATCTAAAGGCC-3') contains the translation start codon and a *Bam*HI site (underlined). The reverse primers for cloning ELG3 and ELG3/V5-His, 5'-ATATCACGATGCGGCCGCTTATTGTGCTTTCTTGTTCATCACTCC-3' and 5'-ATATCACGATGCGGCCGCTTTTGTGCTTTCTTGTTCATCACTCC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG3 provides the translation stop codon. The reverse primer for cloning ELG3/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Advantage-HF polymerase (Clontech) as per the manufacturer's instructions. cDNA prepared from ZR-75-1 cells (ATCC No. CRL-1500) was used as the

DNA template. This cDNA was prepared by isolating RNA from the ZR-75-1 cells using Trizol reagent (Gibco BRL) as per the manufacturer's instructions and then reverse transcribing the RNA using MuLV reverse transcriptase and random hexamers as described for the GeneAmp RNA PCR kit (PE Applied Biosystems).

PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transform *E. coli* strain TOP10F' (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG3 and ELG3/V5-His were designated pTh1015.1 and pTh1017.1, respectively.

ELG3 was also cloned into the pBEVY-L yeast expression vector (Miller et al., 1998, *Nucl. Acids Res.*, 26: 3577-3583) under the control of the constitutive glyceraldehyde 3-phosphate dehydrogenase promoter. The ELG3 coding sequence was obtained by restricting pTh1015.1 with *Bam*HI and *Xba*I, and gel purifying the ~0.9 kb fragment. The pBEVY vector was restricted with *Bam*HI and *Eco*RI, or *Xba*I and *Eco*RI, and the ~1 kb and ~ 6 kb fragments, respectively, were gel purified. The three fragments were ligated and the ligation products were used to transform *E. coli* strain INVαF' (Invitrogen). A plasmid containing the ELG3 gene was isolated and identified by restriction analysis. The insert DNA was confirmed by DNA sequencing and the plasmid designated pLh5015.1 (Figure 19).

Example 4 - Cloning ELG4

Obtaining Complete Coding Sequence for ELG4

A cDNA clone with an incomplete coding sequence for ELG4 was obtained from the SuperScript human leukocyte cDNA library (Gibco BRL) using the GeneTrapper cDNA Positive Selection System (Gibco BRL) as per the manufacturer's instructions. The sequence of the oligonucleotide used to probe the library and repair the captured cDNA target was 5'-GCCAGCCTACCAGAAGTATTTG-3'. The repaired DNA was used to transform UltraMax DH5α-FT cells (Gibco BRL). A clone containing ELG4 was identified by colony PCR using 5'-GCGCAAGAAAAATAGCCAAG-3' and 5'-AATGATGCACGCAAAGACTG-3' as gene specific primers and Platinum Taq DNA polymerase (Gibco BRL). Plasmid DNA was isolated and the insert was sequenced. The plasmid was designated pSh1026.1. The complete coding sequence for ELG4 could not be determined, however, an open reading frame containing the C-terminus of the ELG4 protein was identified. Subsequent cloning (described below) revealed that pSh1026.1 contains an ELG4 variant with an internal deletion of nucleotides 210-255 of the coding sequence.

The nucleotide sequence obtained from pSh1026.1 was used to design a forward (5'-CACGCGGGATCCCTGATGAATACAGAGCCGTGG-3') and reverse (5'-ATATCACGATGCGGCCGCTCAATTATCTTTGTTTTTGCAAGTTCC-3') primer for cloning ELG4 by PCR. These primers contain a *Bam*HI and *Not*I site, respectively (underlined). The forward primer includes the first possible translation start codon identified in pSh1026.1. The reverse primer provides the translation stop codon.

PCR was carried out using Advantage HF polymerase (Clontech) as per the manufacturer's instructions. The SuperScript human leukocyte cDNA library (Gibco BRL) was used as the DNA template.

The PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT (Invitrogen) cut with the same enzymes. The ligation products were used to transform *E. coli* strain TOP10 (Invitrogen). Plasmids were isolated and their inserts were sequenced. A plasmid containing the complete coding sequence for ELG4 as well as 108 nucleotides of 5'-UTR was designated pTh1030.1.

Cloning ELG4 into Expression Vector

ELG4 was cloned into the pYES2/CT yeast expression vector using PCR. Two plasmid constructions were made for the production of the ELG4 protein with either a C-terminal tag containing the V-5 epitope and polyhistidine peptide (ELG4/V5-His), or the ELG4 protein without the tag (ELG4). The forward primer (5'-CACGCGGGATCCCTGATGGAAAAGCCCATTAATATTC-3') contains the translation start codon and a *Bam*HI site (underlined). The reverse primers for cloning ELG4 and ELG4/V5-His, 5'- ATATCACGATGCGGCCGCTCAATTATCTTTGTTTTTGCAAGTTCC-3' and 5'- ATATCACGATGCGGCCGCCAATTATCTTTGTTTTTGCAAGTTCC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG4 provides the translation stop codon. The reverse primer for cloning ELG4/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Advantage-HF polymerase (Clontech) as per the manufacturer's instructions. pTh1030.1 was used as the DNA template for ELG4 and pTh1021.1 (described below) was used as the template for ELG4/V5-His.

The PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transform *E. coli* strain TOP10 (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG4 and ELG4/V5-His were designated pTh1021.1 and pTh1021.2, respectively.

Example 5 - Cloning ELG5

ELG5 was cloned into the pYES2/CT yeast expression vector (Invitrogen) using PCR. Two plasmid constructions were made for the production of the ELG5 protein with either a C-terminal tag containing the V-5 epitope and polyhistidine peptide (ELG5/V5-His), or the ELG5 protein without the tag (ELG5). The forward primer (5'-CACGCGGGATCCAAAAATGAACATGTCAGTGTTGACTTTACAAG-3') contains the translation start codon and a *Bam*HI site (underlined). The reverse primers for cloning ELG5 and ELG5/V5-His, 5'-ATATCACGATGCGGCCGCCTATTCAGCTTTCGTTGTTTTCCTC-3' and 5'-ATATCACGATGCGGCCGCCATTCAGCTTTCGTTGTTTTCCTC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG5 provides the translation stop codon. The reverse primer for cloning ELG5/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Advantage-HF polymerase (Clontech) as per the manufacturer's instructions. The ProQuest human liver cDNA library (Gibco BRL) was used as the DNA template.

The PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transformed *E. coli* strain TOP10 (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG5 and ELG5/V5-His were designated pTh1018.1 and pTh1019.1, respectively.

Example 6 - Cloning ELG6

ELG6 was cloned into the pYES2/CT yeast expression vector (Invitrogen) using PCR. Two plasmid constructions were made for the production of the ELG6 protein with either a C-terminal tag containing the V-5 epitope and polyhistidine peptide (ELG6/V5-His), or the ELG6 protein without the tag (ELG6). The forward primer (5'-CACGCGGGATCCAAAAATGGTCACAGCCATGAATGTCTC-3') contains the translation start codon and a *Bam*HI site (underlined). The reverse primers for cloning ELG6 and

ELG6/V5-His, 5'-ATATCACGAT<u>GCGGCCGC</u>TCACTGGCTCTTGGTCTTGGC-3' and 5'-ATATCACGAT<u>GCGGCCGC</u>CACTGGCTCTTGGTCTTGGC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG6 provides the translation stop codon. The reverse primer for cloning ELG6/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Advantage-HF polymerase (Clontech) as per the manufacturer's instructions. The SuperScript human leukocyte cDNA library (Gibco BRL) was used as the DNA template.

The PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transform *E. coli* strain TOP10 (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG6 and ELG6/V5-His were designated pTh1041.1 and pTh1042.1, respectively.

Example 7 - Cloning ELG7

ELG7 was cloned into the pYES2/CT yeast expression vector (Invitrogen) using PCR. Two plasmid constructions were made for the production of the ELG7 protein with either a C-terminal tag containing the V-5 epitope and polyhistidine peptide (ELG7/V5-His), or the ELG7 protein without the tag (ELG7). The forward primer (5'-CACGCG<u>GGATCC</u>AAAAATGGGGCTCCTGGACTCGGAGC-3') contains the translation start codon and a *Bam*HI site (underlined). The reverse primers for cloning ELG7 and ELG7/V5-His, 5'-ATATCACGAT<u>GCGGCCGC</u>TTAATCTCCTTTTGCTTTTCCATTTTTCTGC-3' and 5'-ATATCACGAT<u>GCGGCCGC</u>TTATCTCCTTTTGCTTTTCCATTTTTCTGC-3', respectively, contain a *Not*I site (underlined). The reverse primer for cloning ELG7 provides the translation stop codon. The reverse primer for cloning ELG7/V5-His only contains 2 of the 3 bases of the stop codon, therefore, placing the gene in frame with the tag provided by the vector.

PCR was carried out using Platinum Taq DNA polymerase (Gibco BRL) as per the manufacturer's instructions. The SuperScript human leukocyte cDNA library (Gibco BRL) was used as the DNA template.

The PCR products were gel purified, digested with *Bam*HI and *Not*I, and ligated into pYES2/CT cut with the same enzymes. The ligation products were used to transform *E. coli*

strain TOP10 (Invitrogen). Plasmids were isolated and their inserts were sequenced. Plasmids coding for ELG7 and ELG7/V5-His were designated pTh1044.1 and pTh1045.1, respectively.

Example 8 - Determination of Tissue Distribution by Northern Blot Analysis

A membrane containing poly(A)$^+$ RNA from 12 different human tissues (brain, heart, skeletal muscle, colon, thymus, spleen, kidney, liver, small intestine, placenta, lung and peripheral blood leukocytes) was purchased from Clontech (Human 12-lane MTN blot). Northern blot analysis was carried out using standard procedures (Ausubel et al., 1994-, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY). The hybridization solution contained 10% dextran sulphate. Probes were prepared by labelling cDNA using [alpha-^{32}P]dCTP and Rediprime II Random Prime Labelling System (Amersham Pharmacia Biotech). The cDNA probes for ELG1, ELG3, ELG5, and ELG6 corresponded to the complete CDS for the genes. The cDNA probes for ELG2, ELG4 and ELG7 corresponded to bases 209-514, 408-726 and 113-566 of the CDS, respectively. The membrane was washed at high stringency using 0.25X SSC, 0.1% SDS at 55°C. The Northern blots are shown in Figure 27.

Example 9 - Cloning Human ELG1 Control Region

The ELG1 control region (989 bp) is cloned from human leukocyte genomic DNA by PCR. The control region is amplified by PCR using synthetic forward and reverse primers starting at positions -2865 bp and -1877 bp upstream from the translation initiation codon, ATG. The forward and reverse primers used for cloning human ELG1 control region by PCR amplification are 5'-GGA<u>AGATCT</u>TACAGGCTCGTGAGGCTTCCCTCCCG-3' and 5'-GGA<u>AGATCT</u>CCGGCAGGAGGGACCAAGGCT-3', respectively. The *Bgl*II recognition sequence (underlined) is included to facilitate cloning.

The PCR amplification is performed in a Perkin-Elmer GeneAMP PCR system 9700 instrument. For example, the PCR is performed in a 50 µl reaction volume containing 0.5 µg of genomic DNA, 0.4 µM of each primer, 1X dNTP mix (Clontech, CA), 1X cDNA PCR reaction buffer (Clontech) and 1X Advantage cDNA polymerase mix (Clontech).

The conditions for the PCR reaction are:

7 cycles at 94°C for 2 seconds, 72°C for 3 minutes
32 cycles at 94°C for 2 seconds, 67°C for 3 minutes
67°C for 4 minutes

The PCR product is gel-purified using QIAquick gel extraction kit (Qiagen, Germany) and ligated into the TA cloning vector pCRII (Invitrogen) according to manufacturers instruction. The ligation product is used to transform *E. coli* TOP10 strain (Invitrogen). The resulting plasmids are screened by restriction analysis and confirmed by DNA sequencing. The human ELG1 control region is then recloned from the pCRII/ELG1 control region construct into the luciferase reporter vector pGL3-Basic (Promega). The resulting human ELG1 control region/reporter construct is used to transfect different mammalian cell lines, and reporter activity measured.

Example 10 - Cloning Human ELG2 Control Region

The ELG2 control region (509 bp) is cloned from human leukocyte genomic DNA by PCR. The control region is amplified by PCR using synthetic forward and reverse primers starting at positions -53626 bp and -53118 bp upstream from the translation initiation site, ATG. The forward and reverse primers used for cloning human ELG2 control region by PCR amplification are 5′-GGAAGATCTCGAGGGTGGGCTTCTGCCACCC-3′ and 5′-GGAAGATCTCTTTTAGCCCAAGGGGCGGCAGC-3′, respectively. The *Bgl*II recognition sequence (underlined) is included to facilitate cloning.

The PCR amplification and cloning are performed as described in Example 9.

The resulting human ELG2 control region/reporter construct is used to transfect different mammalian cell lines, and reporter activity measured.

Example 11 - Cloning of the Human ELG3 Control Region

The human ELG3 control region was cloned from human leukocyte genomic DNA by nested PCR. Blood was obtained from volunteers in the present inventors' laboratory and used to prepare genomic DNA that served as template. In the first PCR reaction, synthetic forward and reverse primers starting at position -2025 bp and -1 bp, respectively, upstream from the translation initiation codon, ATG of the ELG3 gene were used. The forward and reverse primers were 5′- GGAAGATCTTTCGTGTGAATTTCCTTCAAGTCTC-3′ and 5′-GGAAGATCTTGATCCGCAGCGGCTGTG-3′, respectively. The *Bgl*II recognition sequence (underlined) was included to facilitate cloning.

The PCR amplification was conducted in a Perkin-Elmer GeneAMP PCR system 9700 instrument, in a 50 µl reaction volume containing 0.5 µg of genomic DNA, 0.4 µM of each primer, 1X dNTP mix (Clontech, CA), 1X cDNA PCR reaction buffer (Clontech) and 1X Advantage cDNA polymerase mix (Clontech).

The conditions for the PCR reaction were:

7 cycles at 94°C for 2 seconds, 72°C for 3 minutes
32 cycles at 94°C for 2 seconds, 67°C for 3 minutes
67°C for 4 minutes

Analysis of the PCR product by agarose gel electrophoresis revealed that at least two primer specific bands of about 2 kb were amplified. This result necessitated the use of the PCR products as a template and a new set of internal primers in a second PCR reaction to generate a unique primer specific band corresponding to the ELG3 control region. The internal forward and reverse primers start at positions -1381 and -37 respectively, upstream from the translation initiation codon, ATG. The internal forward and reverse primers used were 5′-GGAAGATCTCCGGTACCTACAGTTACTCACTCTGC-3′ and 5′-GGAAGATCTGGCGATGCGCTGTCCAGGGTA-3′.

The conditions for PCR reaction described herein were used for the second PCR reaction except for the following modifications: the second temperature cycle was lowered from 32 to 22 cycles, Taq DNA polymerase was substituted for cDNA polymerase and Q solution (Qiagen) was used according to manufacturer's instruction.

The PCR product was gel-purified using QIAquick gel extraction kit (Qiagen). The purified PCR product and the reporter vector pGL3-Basic were separately digested with *Bgl*II restriction enzyme to generate compatible ends suitable for in-frame ligation of the PCR product to the luciferase gene of pGL3-basic. The ligation product was used to transform *E. coli* TOP10 strain (Invitrogen). The resulting plasmid, pGh3020.1 (Figure 20), was screened by restriction analysis and confirmed by DNA sequencing. The resulting human ELG3 control region/reporter construct is used to transfect different mammalian cell lines, and reporter activity measured.

<u>Example 12 - Cloning Human ELG4 Control Region</u>

The ELG4 control region (2456 bp) is cloned from human leukocyte genomic DNA by PCR. The control region is amplified by PCR using synthetic forward and reverse primers. The forward and reverse primers used for cloning human ELG4 control region by PCR amplification are 5'-CG<u>ACGCGT</u>TGCGCCTGGCTGAACACTAC-3' and 5'-GGA<u>AGATCT</u>CTGGGACAAACAACAGGC-3', respectively. The *Mlu*I and *Bgl*II recognition sequences (underlined), respectively, are included to facilitate cloning.

The PCR amplification and cloning are performed as described in Example 9.

The resulting human ELG4 control region/reporter construct is used and to transfect different mammalian cell lines, and reporter activity measured.

<u>Example 13 - Cloning Human ELG5 Control Region</u>

The ELG5 control region (1411 bp) is cloned from human leukocyte genomic DNA by PCR. The control region is amplified by PCR using synthetic forward and reverse primers starting at positions -1411 bp and -1 bp upstream the translation initiation codon, ATG. The forward and reverse primers used for cloning human ELG5 control region by PCR amplification are 5'-CCG<u>CTCGAG</u>GTGAGCCACCACCGCGGCC-3' and 5'-CCG<u>CTCGAG</u>TGGGGCTGATCTTCGGAGTCGC-3', respectively. The *Xho*I recognition sequence (underlined) is included to facilitate cloning.

The PCR amplification and cloning are performed as described in Example 9.

The resulting human ELG5 control region/reporter construct is used to transfect different mammalian cell lines, and reporter activity measured.

<u>Example 14 - Cloning Human ELG6 Control Region</u>

The ELG6 control region (1937 bp) is cloned from human leukocyte genomic DNA by PCR. The control region is amplified by PCR using synthetic forward and reverse primers starting at positions -1937 bp and -1 bp upstream the initiation codon, ATG. The forward and reverse primers used for cloning human ELG6 control region by PCR amplification are 5'-

CC<u>GAGCTC</u>GATTAGCTGTCAGGCTATATATGGAGCC-3' and 5'-CC<u>GAGCTC</u>CTAGTTTGCAGAAGGTCCAAAGC-3', respectively. The *Sac*I recognition sequence (underlined) is included to facilitate cloning.

The PCR amplification and cloning are performed as described in Example 9.

The resulting human ELG6 control region/reporter construct is used to transfect different mammalian cell lines, and reporter activity measured.

<u>Example 15 - Cloning Human ELG7 Control Region</u>

The ELG7 control region (2000 bp) is cloned from human leukocyte genomic DNA by PCR. The control region is amplified by PCR using synthetic forward and reverse primers starting at positions -2000 bp and -1 bp upstream the translation initiation codon, ATG. The forward and reverse primers used for cloning human ELG7 control region by PCR amplification are 5'-CC<u>GAGCTC</u>GGAAATACCTGAAGCTGTTTTAAC-3' and 5'-CC<u>GAGCTC</u>CGCGGCGATGAGCGGGC-3', respectively. The *Sac*I recognition sequence (underlined) is included to facilitate cloning.

The PCR amplification and cloning are performed as described in Example 9.

The resulting human ELG7 control region/reporter construct is used to transfect different mammalian cell lines, and reporter activity measured.

<u>Example 16 - Drug Screening Assay Using ELG3 Control Region.</u>

Plasmid pGh3020.1 (Figure 20), containing the ELG3 control region, is used to screen test compounds that modulate the ELG3 promoter activity. Transient transfections are performed to evaluate the functionality of the ELG3 control region using techniques known by persons skilled in the art.

Alternatively, HepG2 cells are stably transfected with 10 μg of pGh3020.1 and 1 μg of vector pRSV-NEO (ATCC), using 10 μl of Lipofectamine 2000 Reagent (Gibco BRL) in a 60 mm tissue culture dish as described by the manufacturer. After a 24 h incubation, the cells are passaged into two 150 mm tissue culture dishes at a 1:2 dilution and grown for another 24 h. Geneticin (Gibco BRL) is added to the medium at a concentration of 800 μg/ml. After 3-4

weeks of growth under the selection pressure of the antibiotic, the resistant clones are isolated and characterized for their luciferase activity.

Drug screening is performed using the Luciferase Enzyme Assay System (Promega), following the manufacturer's recommendations. Briefly, transfected cells grown in a 96 well plate are exposed to test compounds. After an appropriate incubation time, the cells are washed with Mg^{2+} and Ca^{2+} free PBS. Cells are lysed with 20 µl of 1X Luciferase Cell Culture Lysis Reagent (CCLR, Promega). The plate is placed into a luminometer with an automatic injector. For each well, the injector adds 100 µl of Luciferase Assay Reagent (Promega), and the light emission generated by the reaction is read for 10 seconds after a 2 second delay. Cell cultures without a test compound are used as controls. Any significant difference in the luciferase activity indicates that the test compound is modulating the ELG3 promoter activity.

This assay or other reporter assays are suitable for drug screening using the control region of any elongase gene.

Example 17 - Drug Screening Assays Using Yeast One-Hybrid Systems

Methods for yeast one-hybrid assays are known by persons skilled in the art (Fields S. and Song O., 1989, *Nature*, 340: 245-246 and Ulmasov et al., 1997, *Science*, 276: 1865-1868). Reagents and/or kits are commercially available for the assays, e.g., the Matchmaker One-Hybrid System (Clontech).

This assay is suitable for all of the elongase control regions described herein.

The known target elements, or elongase control region 'bait' is inserted upstream of a reporter gene (e.g. *HIS3*) and integrated into the yeast genome to make a new reporter strain. The yeast strain is transformed with an activation domain (AD) fusion library to screen for DNA binding proteins that interact with the bait DNA sequence. Binding of an AD/DNA-binding domain (DBD) hybrid protein to the target sequence results in activation of the reporter gene transcription and subsequent selection. For example, expression of *HIS3* will allow colony growth on minimal medium lacking histidine. The cDNA encoding DNA binding protein (DBP) is isolated and characterized. The interaction is reconstructed *in vitro* or *in vivo* for screening test compounds by exposing the target elements or elongase control region to the DBP in the presence of test compounds. The effect of the test compound is evaluated through assays, well known to those skilled in the art, that measure DNA/protein binding interactions.

Example 18 - Drug Screening Assays Using Yeast Two-Hybrid Systems

Methods for the yeast two-hybrid assays are known by persons skilled in the art (Fields S. and Song O., 1989, *Nature*, 340: 245-246 and Furuyama K. and Sassa S., 2000, *J. Clin. Invest.*, 105: 757-764). Reagents and/or kits are commercially available for the assays, e.g., the Hybrid Hunter Yeast Two-Hybrid (Invitrogen), the Matchmaker Two-Hybrid Systems (Clontech) and the HybriZAP Two Hybrid System (Stratagene).

This assay is suitable for all of the elongase genes disclosed herein.

Two physically distinct functional domains are necessary: a DNA binding domain (DBD) and an activation domain (AD). The elongase polypeptide of interest is cloned into a "bait" vector, and expressed as a hybrid protein with a DBD. A library of cDNAs encoding potential interacting proteins is cloned in frame with AD in the "prey" vector. The bait and prey vector fusion constructs are transformed into one of several engineered yeast strains. If an interaction between bait and prey hybrid proteins occurs, the AD of the prey is brought into close contact with the DBD and transcription of the reporter genes is activated. Positive interacting proteins are easily identified by plating on nutrient deficient medium, and screening for reporter activity.

The interaction between these two proteins is reconstructed *in vitro* or *in vivo* for screening test compounds by exposing the two interacting proteins to test compounds. The effect of the test compound is evaluated through assays, well known to those skilled in the art, that measure protein/protein binding interactions.

Example 19 - Functional Analysis of Human Elongases in *Saccharomyces cerevisiae*

The example presented herein demonstrates that the human elongase genes, ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7 cloned by the inventors, encode enzymes able to elongate, by at least two carbons, n-3 and/or n-6 fatty acid substrates.

Materials

Lithium [1-^{14}C]18:3n-6, [1-^{14}C]18:3n-3, [1-^{14}C]20:4n-6, and [1-^{14}C]20:5n-3 (99% radiochemical purity; specific activity: 48 to 58 µCi/µmol), were purchased from NEN (Boston, MA). All unsaturated fatty acids were saponified with 0.1 M LiOH and dissolved in a synthetic minimal medium lacking uracil (SC-U) with 1% tergitol.

Fatty acid free bovine serum albumin, tergitol, Tris-HCl, carbohydrates, amino acids and fatty acids were obtained from Sigma-Aldrich Canada (ON, Canada). Yeast nitrogen base without amino acids was purchased from Difco (Becton Dickinson). All organic solvents (HPLC grade) were obtained from Fisher-Scientific (Fair Lawn, NJ).

Yeast Transformation

Saccharomyces cerevisiae strain INVSc1 (Invitrogen) was transformed with the elongase constructs previously described (Examples 1-7) or pYES2/CT using the lithium acetate method as supplied by Invitrogen. For the expression of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7 the yeast were transformed with pTh1009.1, pTh1014.1, pTh1015.1, pTh1021.1, pTh1018.1, pTh1041.1 or pTh1044.1, respectively. For the expression of ELG1/V5-His, ELG2/V5-His, ELG3/V5-His, ELG4/V5-His, ELG5/V5-His, ELG6/V5-His or ELG7/V5-His the yeast were transformed with pTh1009.2, pTh1014.2, pTh1017.1, pTh1021.2, pTh1019.1, pTh1042.1 or pTh1045.1, respectively. Recombinant yeast cells were selected on SC-U medium.

Incubation

Transformed yeast (approximately 3.2 x 10^6 cells/ml; O.D.$_{600}$ 0.4) were incubated in a 125 ml Erlenmeyer containing 10 ml of SC-U medium with 1% raffinose, 1% tergitol and 25 μM of the lithium salts of either [1-^{14}C]18:3n-3 (1 μCi), [1-^{14}C]18:3n-6 (1 μCi), [1-^{14}C]20:4n-6 (2 μCi), or [1-^{14}C]20:5n-3 (2 μCi). After 4 h incubation in an orbital incubator at 270 rpm and 30°C, cells reached the log phase and the transgene expression was induced with galactose (2% final concentration). The yeast were incubated for an additional 19 h and then harvested by centrifugation at 5000 x g for 10 minutes at 4°C.

Cells were washed with Tris-HCl buffer (100 mM, pH 8.0) containing 0.1% BSA and total lipids were extracted as described below. The radioactivity from aliquots of the incubation medium, supernatant and cells was determined by liquid scintillation counting using a LS6500-Scintillation System (Beckman).

The host yeast transformed with pYES2/CT was used as negative control.

Lipid Extraction

Total lipids were extracted from cells with chloroform/methanol (2:1 v/v) according to the method of Folch et al., 1957, *J. Biol. Chem.*, 226: 497-509. Alternatively, cells were

resuspended in 1.5 ml of water and saponified with 2 ml of 10% KOH in ethanol. The total lipid extracts or the free fatty acids from the saponified samples were methylated using boron trifluoride in methanol at 90°C for 30 min. The resultant methyl esters (FAME) were analyzed as described below.

Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) Analysis

Analyses of radiolabelled FAME were carried out on a Hewlett Packard 1090, series II chromatograph equipped with a diode array detector set at 205 nm, a radioisotope detector (model 171, Beckman, CA) with a solid scintillation cartridge (97% efficiency for ^{14}C-detection) and a reverse-phase ODS (C-18) Beckman column (250 mm x 4.6 mm i.d.; 5 μm particle size) attached to a pre-column with a μBondapak C-18 (Beckman) insert. FAME were separated isocratically with acetonitrile/water (95:5 v/v) at a flow rate of 1 ml/min and were identified by comparison with authentic standards. Alternatively, the eluted FAME were collected and the solvent evaporated. FAME were re-dissolved in hexane for further analysis by gas chromatography.

Gas Chromatography (GC) Analysis

The FAME profile was determined using a Hewlett Packard Gas Chromatograph equipped with an interfaced ChemStation, a flame-ionization detector and a 30 m x 0.25 mm i.d. fused silica column (HP-wax, cross linked polyethylene glycol, film thickness 0.25 μm) and He as gas carrier. The temperatures of the injector and detector were maintained at 225°C and 250°C, respectively. After an initial hold of 1 min at 180°C, the column temperature was increased by 4°C/min to 190°C (7 min hold), then by 10°C/min to 200°C (5 min hold) and finally by 25°C/min to 215°C. This temperature was maintained for 17.9 min. FAME were identified by comparison with authentic standards.

Results

RP-HPLC analyses revealed that the exogenously added radiolabelled polyunsaturated fatty acids were elongated by at least two carbons in yeast transformed with human elongase genes (Table 3). In yeast expressing ELG4, 18:3n-6 was converted into 20:3n-6 which was then elongated to 22:3n-6, 20:4n-6 was converted into 22:4n-6 which was further elongated to 24:4n-6 and 18:3n-3 was converted into 20:3n-3 and 22:3n-3 (Figure 21). Yeast transformed with pYES2/CT did not elongate any of these substrates (Figure 22).

In yeast expressing elongases with V5-His tag, the percent elongation of selected substrates was similar to that detected in yeast with non-tagged enzymes (Table 4).

Conclusion

The functional analysis of the human ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7 genes confirmed that each gene encodes a fatty acid elongase which is active on various PUFAs.

Table 3

Percent Elongation of PUFA Substrates to their Products in Yeast Expressing Human Elongases

Gene	Plasmid	18:3n-6		20:4n-6		18:3n-3		20:5n-3	
		20:3	22:3	22:4	24:4	20:3	22:3	22:5	24:5
ELG1	pTh1009.1	2	nd	6	2	1	nd	2	nd
ELG2	pTh1014.1	62	3	39	1	16	nd	59	nd
ELG3	pTh1015.1	10	nd	11	21	2	nd	16	29
ELG4	pTh1021.1	20	4	24	2	10	4	15	3
ELG5	pTh1018.1	3	nd	nd	nd	9	nd	-	-
ELG6	pTh1041.1	2	nd	nd	nd	3	nd	nd	nd
ELG7	pTh1044.1	nd	nd	nd	nd	5	nd	nd	nd

nd: not detected
-: not tested

Table 4

Percent Elongation of PUFA Substrates to their Products in Yeast Expressing V5-His Tagged Human Elongases

Gene	Plasmid	18:3n-6		20:4n-6		18:3n-3		20:5n-3	
		20:3	22:3	22:4	24:4	20:3	22:3	22:5	24:5
ELG1	pTh1009.2	-	-	7	nd	-	-	-	-
ELG2	pTh1014.2	73	11	-	-	-	-	-	-
ELG3	pTh1017.1	-	-	8	15	-	-	-	-
ELG4	pTh1021.2	-	-	12	nd	-	-	-	-
ELG5	pTh1019.1	5	-	-	-	-	-	-	-
ELG6	pTh1042.1	nd	nd.	Nd	nd	3	nd	nd	nd
ELG7	pTh1045.1	nd	nd	Nd	nd	4	nd	nd	nd

nd: not detected
-: not tested

Example 20 - Drug Screening Assay for Elongases Using Yeast

This example provides a methodology suitable for screening test compounds that modulate the activity of recombinant elongases in whole cells and spheroplasts of *Saccharomyces cerevisiae*. The test compound uptake is likely to be enhanced in yeast spheroplasts due to their lack of a cell wall. Thus, this is the model of choice for assessing the effect of low concentrations of test compounds on elongase activity.

Spheroplast Preparation

Saccharomyces cerevisiae heterologous for any of the human elongase genes are grown in SC-U medium with 1% raffinose and 2% galactose to induce the expression of the transgene. After 16 h incubation, cells are centrifuged at 2060 x g for 5 min at 4°C, washed once with distilled water and centrifuged again. The volume and weight of the cell pellet are measured. Cells are suspended (1:2 w/v) in 0.1 M Tris.SO_4 (pH 9.4), 10 mM DTT and incubated at 30°C.

After 10 min incubation, the cell pellet is obtained by centrifugation, washed once (1:20 w/v) with 1.2 M sorbitol and suspended (1:1 w/v) in 1.2 M sorbitol, 20 mM phosphate buffer (pH 7.4) as described elsewhere (Daum et al., 1982, *J. Biol. Chem.*, 257: 13028-13033). A 15,800 x g (1 min) supernatant of lyticase is added to the cell suspension at a concentration of 2000 U/ml and the suspension incubated at 30°C with 50 rpm shaking. Conversion to spheroplasts is checked after 40 min incubation by diluting the suspension with distilled water followed by observation under the microscope (Schatz G. and Kovac L., 1974, *Meth. Enzymol.*, 31A: 627-632). After 70 min incubation, approximately 90% of the cells are converted to spheroplasts.

Incubation of Spheroplasts with Test Compounds

Spheroplasts are harvested by centrifugation at 2060 x g for 5 min at 4°C and washed once with 1.2 M sorbitol. Spheroplasts are resuspended in SC-U medium with 1% raffinose, 1% tergitol, 1.2 M sorbitol and 2% galactose to maintain the induction conditions and to give an $O.D._{600}$ reading of approximately 2.5-3.0. A 10 ml aliquot of the spheroplast suspension is transferred to a 125 ml Erlenmeyer flask and incubated with 200 µl of a test compound in ethanol (e.g. pebulate sulphoxide with a final concentration ranging from 0.01 to 100 µM) at 30°C in an orbital incubator at 270 rpm. After 30 min incubation, 1 µCi of a selected elongase substrate (i.e., lithium salts of [1-^{14}C]18:3n-6, [1-^{14}C]20:4n-6, [1-^{14}C]20:5n-3 or [1-^{14}C]18:3n-3) is added to the culture to a final concentration of 2 to 200 µM and further incubated for 120 min. Cell density is determined ($O.D._{600}$) and spheroplasts are harvested by centrifugation and washed with Tris-HCl buffer (100 mM, pH 8.0) containing 0.1% BSA. Total lipids are extracted and analyzed as described in Example 19.

Incubation of Whole Yeast with Test Compounds

Saccharomyces cerevisiae heterologous for any of the human elongase genes are incubated in a 125 ml Erlenmeyer flask containing 9 ml of SC-U medium with 1% raffinose, 1% tergitol ($O.D._{600}$ 0.4, approximately 3.2 x 10^6 cells/ml) and 200 µl of a test compound in ethanol (e.g. pebulate sulphoxide, with a final concentration in the culture that range between 0.1 and 5 mM). After 1 h incubation in an orbital incubator at 270 rpm and 30°C, 1 µCi of a selected elongase substrate (i.e., lithium salts of [1-^{14}C]18:3n-6, [1-^{14}C]20:4n-6, [1-^{14}C]20:5n-3 or [1-^{14}C]18:3n-3) is added to the culture to a final concentration of 2 to 200 µM. After 4 h incubation with the inhibitor, cells reach the log phase and the transgene expression is induced with the addition of 1 ml of galactose to a final concentration of 2%. The yeast are incubated for an additional 19 h and then harvested by centrifugation at 5000 x g for 10 minutes at 4°C. Cells are washed with Tris-HCl buffer (100 mM, pH 8.0) containing 0.1% BSA and total lipids are extracted and analyzed as described in Example 19.

Calculations

The elongase activity is determined by measuring the conversion of radiolabelled 18:3n-6 to 20:3n-6 and 22:3n-6, 20:4n-6 to 22:4n-6 and 24:4n-6, 18:3n-3 to 20:3n-3 and 22:3n-3 or 20:5n-3 to 22:5n-3 and 24:5n-3. The percent inhibition is calculated as described elsewhere (Kawashima et al., 1996, *Biosci. Biotech. Biochem.*, 60: 1672-1676):

% Inhibition = 100(activity without the inhibitor - activity with the inhibitor)/activity without the inhibitor

Example 21 - Drug Screening Assay for Elongase Using Yeast Microsomes

This example teaches that microsomes from yeast with elongase transgenes contain all the enzymes required for testing the effect of test compounds on the activity of a specific recombinant fatty acid elongase.

Materials

A sulphoxide derivative of S-propylbutylethylthiocarbamate (pebulate sulphoxide) was obtained from Zeneca Agrochemicals, UK, and dissolved in ethanol at a concentration of 5 mM.

Yeast Microsome Preparation

A 5 l culture of *Saccharomyces cerevisiae* transformed with pTh1017.1 encoding ELG3/V5-His was started with a cell density of approximately 3.2 x10^5 cells/ml ($O.D._{600}$ 0.4) using SC-U medium with 1% raffinose. After 8 h of incubation at 30°C in an orbital shaker at 270 rpm, galactose was added to a final concentration of 2%. Yeast were incubated for an additional 12 h until they were harvested by centrifugation at 2060 x g for 10 minutes at 4°C and washed with water. The cell pellet was resuspended in 1/3 of its volume in a pH 7.2 isolation buffer (80 mM Hepes-KOH, 10 mM KCl, 320 mM sucrose, 2 mM PMSF and a protease inhibitor cocktail). The cell suspension was poured into a mortar containing liquid N_2 and ground with sand using a ceramic pestle. The yeast powder was transferred to a conical test tube, to which 2/3 of the pellet volume of isolation buffer was added. The sand was removed by centrifugation at 228 x g for 1 min and the suspension centrifuged at 10,000 x g for 20 min to separate cell debris, nuclei and mitochondria. The supernatant was centrifuged at 106,000 x g for 1.5 h to obtain the microsomal pellet, which was resuspended in storage buffer (80 mM Hepes-KOH, 10 mM KCl, 320 mM sucrose, 1 mM PMSF and a protease inhibitor cocktail) to

a final protein concentration of 20 µg/µl. The protein concentration was measured by the method of Lowry et al. (1951, *J. Biol. Chem.*, 193: 265-275) with bovine serum albumin as standard.

Incubation of Yeast Microsomes with Pebulate Sulphoxide

The activity of ELG3/V5-His was determined by measuring the conversion of [1-^{14}C]20:5n-3 to [1-^{14}C]22:5n-3 and [1-^{14}C]24:5n-3. Reactions were started by adding 500 µg of yeast microsomal protein to pre-incubated tubes containing 0.20 µCi of the substrate fatty acid at a final concentration of 7.2 µM in 0.25 ml of 80 mM Hepes-KOH (pH 7.2) with 43 mM $MgCl_2$, 1.0 mM ATP, 500 µM NADPH, 10 µM coenzyme A, 100 µM malonyl-CoA (as lithium salt) and pebulate sulphoxide at concentrations that ranged between 1 to 100 µM. The tubes were vortexed vigorously and after 30 min incubation at 37°C in a shaking water bath, the reactions were stopped by the addition of 2 ml of 10% (w/v) KOH in ethanol. Lipids in the incubation mixture were saponified at 80°C for 45 min under N_2. The samples were then left in ice for 5 min before acidification with 750 µl of concentrated HCl. The fatty acids were extracted with hexane and esterified with BF_3 in methanol at 90°C for 30 min. The fatty acid methyl esters were analyzed by HPLC as described in Example 19.

Results

The enzyme activity was expressed in percent conversion of radiolabelled 20:5n-3 into its elongation products. Alternatively, it can be expressed in pmol of the fatty acids produced/mg microsomal protein/min.

Table 5 shows the effect of a thiocarbamate derivative (pebulate sulphoxide) on the ELG3/V5-His activity when 20:5n-3 was provided as substrate. Pebulate sulphoxide at 100 µM substantially reduced elongation, by approximately 27%. This effect was mainly due to a reduction in the synthesis of 22:5n-3 rather than in the production of its metabolite, 24:5n-3.

Table 5

Effect of Pebulate Sulphoxide on the Elongation of [1-^{14}C]20:5n-3 in Microsomes of Yeast Expressing ELG3/V5-His.

Pebulate sulphoxide [µM]	% conversion		
	22:5n-3	24:5n-3	Total
0	13.7	5.0	18.7
1	13.8	5.6	19.4
10	12.8	6.6	19.4
50	11.3	4.6	15.9
100	9.4	4.3	13.7

Values expressed are the average (dispersion ≤ 10%) of two determinations.

Example 22 - Isolation of Recombinant Elongases from Yeast.

This example provides a methodology for the isolation of recombinant elongase from yeast homogenate or microsomes. The purified enzyme is useful for drug screening or for antibody production.

Yeast Homogenate and Microsome Preparation

Yeast cell fractionation was performed as described in Example 21 using yeast expressing ELG3/V5-His.

Elongase Solubilization

Yeast cell homogenate or yeast microsomes were resuspended in solubilization buffer (80 mM HEPES-KOH pH 7.2, 10 mM KCl, 320 mM sucrose, 1 mM PMSF, protease inhibitor cocktail, and 0.5 M NaCl) at 1.3 or 4 mg/ml, respectively. Zwittergent 3-14, *n*-octyl-beta-glucopyranoside or *n*-octyl-beta-thioglucopyranoside (Calbiochem, CA) was added to a final concentration of 2%, with a detergent:protein ratio of 15:1. The mixture was incubated for 2 h at 4°C with stirring and then centrifuged at 106,000 x g for 1 h. The supernatant was removed and stored at -80°C until use. The pellet was resuspended in 1/4 volume of the supernatant using solubilization buffer. The efficiency of each detergent to solubilize the elongase was determined by Western blot analysis as described below.

SDS-PAGE and Western Blot Analysis

Supernant (60 μl) or pellet suspension (20 μl) was mixed with 15 μl or 5 μl of 5X sample loading buffer (1X concentration: 50 mM Tris-HCl pH 8.0, 2% SDS, 10 mM beta-mercaptoethanol, 0.1% bromophenol blue, 10% glycerol), respectively, and boiled at 100°C for 5 minutes. Molecular weight standards (Santa Cruz Biotechnology, CA), controls, 25 μl of the supernatant, and 12.5 μl of the pellet were loaded on 12% pre-cast SDS-polyacrylamide gels. After electrophoresis, the protein was electro-transferred onto a PVDF membrane (Bio-Rad). The membrane was incubated with a blocking solution and subsequently probed with an anti-V5-HRP antibody as recommended by the manufacturer (Invitrogen). The membrane was washed and the antibody was detected using the enhanced chemiluminescence reagent, ECL (Amersham-Pharmacia Biotech.). The membrane was exposed to autoradiography film (Labscientific, NJ).

Zwittergent 3-14 was the most effective detergent in solubilizing ELG3/V5-His, the majority of the tagged protein having been detected in the 106,000 x g supernatant.

Immobilized Metal Ion Affinity Chromatography (IMAC)

The supernatant containing the solubilized enzyme is loaded onto a pre-equilibrated HiTrap chelating (Ni^{2+} charged iminodiacetate) column (Pharmacia) attached to a fast protein liquid chromatography system (Pharmacia). The column is washed with 50 mM sodium phosphate pH 8.0. The tagged protein is eluted with the same buffer containing imidazole ranging from 0 to 500 mM and further concentrated by ultrafiltration using Centriprep (Amicon) concentrators.

Alternatively, Macro-Prep ceramic hydroxyapatite (Bio-Rad, CA), TALON metal affinity resin (a Cobalt-based IMAC resin, Clontech, CA), Ni-nitriloacetic acid resin (Novagen, WI) or other similar resin is used.

Example 23 - Drug Screening Assay for Elongase Using Purified Enzyme

The concentrated enzyme (Example 22) is incubated at 30-37°C in 0.25 ml of 80 mM Hepes-KOH (pH 7.2) with 6 mM egg phosphatidylcholine, 2% Triton X-100, 0.4% sodium deoxycholate, 43 mM $MgCl_2$, 1.0 mM ATP, 500 μM NADPH, 10 μM coenzyme A, 100 μM malonyl-CoA (as lithium salt), 0.20 μCi of the substrate fatty acid (i.e., radiolabelled eicosapentaenoyl-CoA) at a final concentration of 7.2 μM and a test compound (e.g., pebulate

sulphoxide) at concentrations ranging between 0.01 to 100 μM. The tubes are vortexed vigorously and after 30 min incubation at 37°C in a shaking water bath the reactions are stopped by the addition of 2 ml of 10% (w/v) KOH in ethanol.

Total lipids are extracted and methyl ester analyzed as described in Example 19.

Example 24 - Validation of Drug Screening Assays Described in Examples 20, 21 and 23 Using Rat Liver Microsomes

Preparation of Rat Liver Microsomes

Wistar rats under light halothane (15% in mineral oil) anesthesia were sacrificed by exsanguination during periods of high enzyme activity. Livers were immediately rinsed with cold 0.9% NaCl solution, weighed and minced with scissors. All procedures were performed at 4°C unless specified otherwise. Livers were homogenized in a solution (1:3 w/v) containing 0.25 M sucrose, 62 mM potassium phosphate buffer (pH 7.0), 0.15 M KCl, 1.5 mM *N*-acetylcysteine, 5 mM $MgCl_2$, and 0.1 mM EDTA using 4 strokes of a Potter-Elvehjem tissue homogenizer. The homogenate was centrifuged at 10,400 x g for 20 min to pellet mitochondria and cellular debris. The supernatant was filtered through a 3-layer cheesecloth and centrifuged at 105,000 x g for 60 min. The microsomal pellet was gently resuspended in the same homogenization solution with a small glass/teflon homogenizer and stored at -80°C. The absence of mitochondrial contamination was enzymatically assessed as described elsewhere (Kilberg, M.S. and Christensen H.N., 1979, *Biochemistry*, 18: 1525-1530). The protein concentration was measured by the method of Lowry et al (1951, *J. Biol. Chem.*, 193: 265-275) with bovine serum albumin as standard.

Incubation of Rat Liver Microsomes with Test Compounds

Reactions were performed using 500 μg of rat liver microsomal protein with the same concentrations of pebulate sulphoxide, radiolabelled fatty acid, conditions and procedures described in Example 21.

Results

The enzyme activity was expressed in percent conversion of radiolabelled 20:5n-3 into its elongation and final delta-6-desaturation products (i.e., 22:5n-3, 24:5n-3 and 24:6n-3). When the incubation was performed under nitrogen, the desaturation reaction did not occur.

Table 6 shows the effect of a thiocarbamate derivative (pebulate sulphoxide) on the rat liver elongase activity when 20:5n-3 was provided as substrate. Pebulate sulphoxide (100 μM)

reduced elongation by approximately 30%. This effect was mainly due to a reduction in the synthesis of 24:5n-3 rather than in the synthesis of 22:5n-3.

Table 6

Effect of Pebulate Sulphoxide on the Elongation of [1-^{14}C]20:5n-3 in Rat Liver Microsomes

Pebulate sulphoxide [μM]	% conversion			
	22:5n-3	24:5n-3	24:6n-3*	Total
0	11.6	39.7	9.1	60.4
1	12.5	47.5	9.6	69.3
10	12.5	47.2	10.9	70.7
50	12.2	48.7	7.9	68.8
100	10.2	28.0	4.5	42.7

Values are expressed as the mean (dispersion ≤ 10%) of two determinations.

* 24:6n-3 is the product of a delta-6-desaturation of 24:5n-3.

Since the rat liver microsomal and the recombinant human elongase (Example 21) activities were similarly affected by pebulate sulphoxide, it is concluded that rat liver microsomes are suitable to use in the validation of drug screening assays.

Example 25 - Functional Characterization of Recombinant Fatty Acid Elongase and Desaturase in Yeast Co-expressing ELG3 and D6D

This example shows a partial reconstitution of the n-3 and n-6 polyunsaturated fatty acid biosynthetic pathway in a heterologous host such as *Saccharomyces cerevisiae* using human fatty acid elongase and desaturase genes .

Materials

[1-^{14}C]18:3n-3, [1-^{14}C]20:4n-6, [1-^{14}C]20:5n-3 and [1-^{14}C]18:2n-6 (99% radiochemical purity; specific activity: 51 to 56 μCi/μmol) were purchased from NEN (Boston, MA). Fatty acids were saponified with 0.1 M LiOH and dissolved in synthetic minimal medium lacking either leucine (SC-Leu) or uracil and leucine (SC-U-Leu), containing 1% tergitol.

Yeast Transformation

Saccharomyces cerevisiae strain INVSc1 (Invitrogen) was transformed using the lithium acetate method as supplied by Invitrogen. The coding sequence for human delta-6-desaturase (GenBank Accession No. AF126799) was previously cloned into the pYES2/CT vector for the production of the protein with a C-terminal tag containing the V-5 epitope and polyhistidine peptide (D6D/V5-His) as described in Canadian Patent Application No. 2,301,158, Mar., 2000, Winther et al. (plasmid designated pTh5002.1). For the co-expression of ELG3 and D6D/V5-His, the yeast were initially transformed with pTh5002.1. Recombinant yeast cells were selected on SC-U medium and then transformed with pLh5015.1 (Example 3). Double recombinant yeast cells containing both pTh5002.1 and pLh5015.1 were selected on SC-U-Leu medium. Yeast cells transformed with pBEVY-L alone, the cloning vector for ELG3, were selected on SC-Leu medium.

Incubation

Transformed yeast cultures (approximately 3.2 x 10^6 cells/ml; $O.D._{600}$ 0.4) were divided in two experimental groups. The first group was incubated in a 125 ml Erlenmeyer flask containing 10 ml of SC-U-Leu medium with 2% raffinose, 1% tergitol and 25 μM lithium [1-^{14}C]20:4n-6 (1μCi). Yeast of the second group were incubated in 10 ml of SC-U-Leu medium containing 1% raffinose, 2% galactose (to induce the expression of D6D/V5-His) and 1% tergitol. Lithium salts (1 μCi) of either [1-^{14}C]18:3n-3, [1-^{14}C]20:4n-6, [1-^{14}C]20:5n-3 or [1-^{14}C]18:2n-6 were added to both experimental groups at a final concentration of 25 μM. After 24 h incubation in an orbital incubator at 270 rpm and 30°C, cells were harvested by centrifugation at 5000 x g for 10 minutes at 4°C.

The cell pellet was washed with Tris-HCl buffer (100 mM, pH 8.0) containing 0.1% BSA total lipids were extracted and radiolabelled fatty acids analyzed as described in Example 19.

The host yeast transformed with pBEVY-L was used as negative control.

Results

Figures 23 and 24 show that only elongation products of PUFA substrates for ELG3 were detected when galactose was absent from the culture medium since the expression of D6D/V5-His was not induced. The constitutively expressed ELG3 was able to elongate 20:4n-6 to 22:4n-6 and 24:4n-6, 20:5n-3 to 22:5n-3 and 24:5n-3, and to a lesser extent 18:3n-3 to 20:3n-

3. These findings are consistent with those described in Example 19. ELG3 did not elongate 18:2n-6.

The elongation products of PUFA substrates for ELG3 were desaturated by D6D/V5-His when galactose was added to the medium (Figure 24). In this regard, 24:5n-6 and 24:6n-3 were produced from 24:4n-6 and 24:5n-3, respectively.

In the presence of galactose, transformed yeast were also able to delta-6-desaturate 18:2n-6 and 18:3n-3 to 18:3n-6 and 18:4n-3, respectively. These products were then substrates of the ELG3, which elongated them to 20:3n-6 and 20:4n-3, respectively.

Both ELG3 and D6D/V5-His seemed to be more active on n-3 than on n-6 fatty acid substrates.

Yeast transgenic for the human elongase, ELG3, and a human D6D, were able to generate polyunsaturated fatty acids of the so called "Sprecher pathway" (Sprecher H., 2000, *Biochim. Biophys. Acta*, 1486: 219-231). The present inventors are the first to report that products of human ELG3, 24:4n-6 and 24:5n-3, are substrates of a human D6D, which is also active on 18:2n-6 and 18:3n-3.

Example 26 - Functional Characterization of Recombinant Fatty Acid Elongase and Desaturase in Yeast Co-expressing ELG3 and D5D

This example expands the inventors' findings described in Example 25. The sequential elongation and desaturation of n-3 and n-6 PUFAs in a heterologous host co-expressing human fatty acid elongase and D5D genes is demonstrated.

Materials

[1-^{14}C]18:3n-3, [1-^{14}C]20:3n-6 and [1-^{14}C]18:2n-6 (99% radiochemical purity; specific activity: 50 to 52 µCi/µmol) were purchased from NEN (Boston, MA). [1-^{14}C]-$\Delta^{8,11,14,17}$ eicosatetraenoic acid, 20:4n-3, (99% radiochemical purity; specific activity: 55 µCi/µmol) was purchased from ARC (St Louis, MO). Fatty acids were saponified with 0.1 M LiOH and dissolved in either SC-Leu or SC-U-Leu medium, containing 1% tergitol.

Yeast Transformation

Saccharomyces cerevisiae strain INVSc1 (Invitrogen) was transformed using the lithium acetate method as supplied by Invitrogen. The coding sequence for human delta-5-desaturase (GenBank Accession No. AF199596) was previously cloned into the pYES2/CT vector for the production of the protein with a C-terminal tag containing the V-5 epitope and polyhistidine peptide (D5D/V5-His) as described in Canadian Patent Application No. 2,301,158, Mar., 2000, Winther et al.(plasmid designated pTh5009.1). For the co-expression of ELG3 and D5D/V5-His, the yeast were initially transformed with pTh5009.1. Recombinant yeast cells were selected on SC-U medium and then transformed with pLh5015.1 (described in Example 3). Double recombinant yeast cells containing both pTh5009.1 and pLh5015.1 were selected on SC-U-Leu medium. Yeast cells transformed with pBEVY-L alone, the cloning vector for ELG3, were selected on SC-Leu medium.

Incubation

Cultures of transformed yeast (approximately 3.2 x 10^6 cells/ml; $O.D._{600}$ 0.4) were divided in two experimental groups. In the first group, cells were incubated in a 125 ml Erlenmeyer flask containing 10 ml of SC-U-Leu medium with 2% raffinose and 1% tergitol. In the second group, yeast were incubated in 10 ml of SC-U-Leu medium with 1% raffinose, 2% galactose (to induce the expression of D5D/V5-His) and 1% tergitol. Lithium salts (1 µCi) of either [1-^{14}C]18:3n-3, [1-^{14}C]20:3n-6, [1-^{14}C]18:2n-6, or [1-^{14}C]20:4n-3 were added to both experimental groups at a final concentration of 25 µM. After 24 h incubation in an orbital incubator at 270 rpm and 30°C, cells were harvested by centrifugation at 5000 x g for 10 minutes at 4°C.

The cell pellet was washed with Tris-HCl buffer (100 mM, pH 8.0) containing 0.1% BSA, total lipids were extracted and radiolabelled fatty acids were analyzed as described in Example 19.

The host yeast transformed with pBEVY-L was used as negative control.

Results

Figure 25 shows that 20:3n-6 was desaturated to 20:4n-6, which was further elongated to 22:4n-6 and 24:4n-6, when the yeast co-expressed both genes in the presence of galactose. When galactose was not added to the medium, 20:3n-6 was only elongated to 22:3n-6.

Similarly, D5D/V5-His desaturated 20:4n-3 producing 20:5n-3, which was then elongated to 22:5n-3 and 24:5n-3. The elongation of 20:4n-3 to 22:4n-3 and 24:4n-3 was also detected.

Under these experimental conditions, yeast co-expressing both genes was not able to elongate and further desaturate 18:2n-6. D5D/V5-His was not active on 20:3n-3, the direct elongation product of 18:3n-3 generated by ELG3 (Figure 26).

Conclusion

Yeast co-expressing ELG3 and a human D5D, both cloned by the inventors, were able to generate substrates (i.e., 24:4n-6 and 24:5n-3) of the so called "Sprecher pathway" (Sprecher H., 2000, *Biochim. Biophys. Acta*, 1486: 219-231).

Example 27 - Drug Screening Assays Using Whole Cells, Spheroplasts or Microsomes of Yeast Co-Expressing ELG3 and either Human D6D or D5D

The following assays are designed to identify compounds that affect the human elongase ELG3 and/or the human desaturases using one host system or any part thereof.

Spheroplast and Microsome Preparation

Transformed *Saccharomyces cerevisiae* cells are grown in SC-U-Leu medium with 1% raffinose and 2% galactose to induce the expression of the desaturase transgenes. After 16 h incubation, spheroplasts are obtained as described in Example 20.

Microsomes from host cells expressing both elongase and desaturase genes are prepared using the liquid N_2 and differential centrifugation methods described in Example 21.

Incubation of Whole Yeast Cells, Spheroplasts or Microsomes with Test Compounds

In these assays with yeast cells containing elongase and desaturase transgenes, the use of SC-U-Leu medium is required to maintain selection pressure. Transformed yeast are incubated with or without galactose to asses the effect of the test component on the activity of ELG3 and the desaturases or the elongase alone, respectively. The substrates of choice are 20:3n-6 or 20:5n-6 for yeast expressing ELG3 and D5D or ELG3 and D6D, respectively. The incubation conditions of whole yeast cells, spheroplasts or microsomes with test compounds are the same as those described in Examples 20 and 21. Regardless of the host system used, the effect of the test compound on the activity of the recombinant enzymes is determined by the RP-HPLC or GC analysis of the relative amounts of FAME produced by ELG3 and/or the desaturases as described in Example 19.

Example 28 - Elongation of PUFAs in Primary Cultures of Leukocytes from Control and STZ-Induced Diabetic Rats

The present example describes the capability of leukocytes to elongate but not desaturate PUFAs. The example also provides details of how the elongation of 18:3n-6 and 18:2n-6 is affected in rats with STZ-induced diabetes.

Materials

RPMI 1640 medium was obtained from Gibco BRL. Streptozotocin (2-desoxy-2-methylnitrosoamino carbonyl amino-D-glucopyranose) was supplied by Sigma.

Animals

Female Wistar rats were obtained from Charles Rivers, St-Constant, Quebec. Animals were housed in barrier-maintained rooms at 22±2°C, a target relative humidity of 50±10% with 15 air changes per hour and a 12 h light/dark cycle. Water and regular chow were provided *ad libitum*.

All animals were monitored daily according to standard procedures in compliance with the Canadian Council of Animal Care guidelines for animal experimentation. Fifteen randomly selected rats were intraperitoneally (I.P.) injected with 50 mg of STZ per kg of body weight. Nine days later, animals received a second dose of STZ (75 mg/kg body weight). A second group of 12 rats which were sham injected with sterile 0.9% NaCl served as control. Two and 7 weeks after the last I.P. injection, control and STZ-treated rats (blood glucose levels 21 to >33 mmoles/l) were put under light halothane (15% in mineral oil) anesthesia and sacrificed by exsanguination. Blood was collected into a 10 ml syringe containing 200 µl of a 5% solution of EDTA as anticoagulant.

Leukocyte Isolation

Leukocytes were obtained by mixing 1 volume of whole blood with 5 volumes of sterile erythrocyte lysis buffer (Qiagen, CA). The cell suspension was incubated for 20 min on ice and centrifuged at 400 x g for 10 min at 4°C. The supernatant was discarded and the leukocyte pellet was washed and resuspended in 550 µl of 0.9% saline. Aliquots were taken for cell counting. Cellular protein content was measured using the method of Lowry et al (1951, *J. Biol. Chem.*, 193: 265-275) with bovine serum albumin as standard.

Incubation

The present inventors' preliminary studies carried out with leukocytes isolated from Wistar rats showed that leukocytes can elongate 18:2n-6, 18:3n-3, 18:3n-6, 20:3n-6 and 20:4n-6 with the elongation of 18:2n-6 and 18:3n-6 being 6% and 66%, respectively, within 24 h. Based on these results and due to the impairment of D6D in diabetes, 18:2n-6 and 18:3n-6, substrate and product of D6D, respectively, were selected for the incubation of leukocytes from control and STZ-induced diabetic rats. No delta-6-desaturation on 18:2n-6, 18:3n-3 or delta-5-desaturation on 20:3n-6, was detected.

Leukocytes from the 2 and 7 week control group, as well as from the 2 and 7 week STZ-treated rat group, were incubated in RPMI 1640 medium with glutamine, 10% fetal calf serum and antibiotics (50 IU/ml penicillin, 50 μg/ml streptomycin) with 5 μM [1-^{14}C]18:3n-6 (0.6 μCi) for 10 min to 24 h or with 5 μM of [1-^{14}C]18:2n-6 (0.6 μCi) for 24 h.

At the end of each incubation, the cell pellet was obtained by centrifugation at 400 x g for 10 min at 4°C. Cells were washed with PBS containing 0.1% bovine serum albumin. Total cellular lipids were extracted with chloroform:methanol (2:1 v/v). Fatty acids were methylated with BF_3 and analyzed by RP-HPLC as described in Example 19. Alternatively, FAME can be analyzed by GC as described in Example 19.

Results

Table 7 shows that leukocytes from STZ-induced diabetic rats rapidly converted 18:3n-6 into 20:3n-6. There was a significant increase in the activity of the elongation system in the STZ group, regardless of the time after the last I.P. STZ injection. Conversely, there was an approximately 50% reduction in the elongation of 18:2n-6 to 20:2n-6 in leukocytes obtained 2 weeks after the STZ injection (Table 8). There were no significant changes in the elongation of 18:2n-6 to 20:2n-6 in leukocytes from animals sacrificed 7 weeks after the STZ treatment.

Table 7

Conversion of 18:3n-6 into 20:3n-6 in Leukocytes from STZ-Induced Diabetic Rats Sacrificed 2 or 7 Weeks Post-Induction

Incubation time (h)	2 weeks		7 weeks	
	STZ	Control	STZ	Control
0	0	0	0	0
0.16	50 ± 8	31 ± 9	37 ± 9	33 ± 4
0.5	115 ± 26	70 ± 12	112 ± 10	71 ± 15
1	288 ± 23	200 ± 16	190 ± 92	143 ± 31
24	nt	nt	1008 ± 98	628 ± 156

Values are expressed in pmol of 20:3n-6 produced/mg cellular protein and represent the mean ± S.D. of 6 rats.
nt: not tested

Table 8

Conversion of 18:2n-6 into 20:2n-6 in Leukocytes from STZ-Induced Diabetic Rats Sacrificed 2 or 7 Weeks Post-Induction

2 weeks		7 weeks	
STZ	Control	STZ	Control
322 ± 119	126 ± 27	147 ± 22	128 ± 32

Leukocytes were incubated for 24 h.
Values are expressed in pmol of 20:2n-6 produced/mg cellular protein and represent the mean ± S.D. of 6 rats.

PUFA metabolism is altered in leukocytes of rats with STZ-induced diabetes. Therefore, leukocytes are an appropriate model to assess the modification or regulation of the elongation system in disease (e.g., diabetes).

Example 29 - Elongation of PUFAs in Primary Cultures of Leukocytes from Humans

This example shows that human leukocytes are a suitable model to assess elongase activity on 18:3n-6. This assay may be used in clinical trials to determine alterations in the elongation system in diseases such as diabetes.

Peripheral venous blood from fasted healthy volunteers (30 to 50 years of age) was obtained using 10 ml Vacutainers (Vacutainer Systems, NJ) containing EDTA as anticoagulant. Leukocytes were isolated using the techniques described in Example 28. The incubation of leukocytes with 5 μM [1-^{14}C]18:3n-6 (0.6 μCi) for 10 to 60 min was performed under the same conditions described in Example 28.

Results

Table 9 demonstrates that human leukocytes have a capability to rapidly elongate 18:3n-6 to 20:3n-6, similar to that found in rat leukocytes (Example 28). No delta-5-desaturation activity was detected on 20:3n-6.

Table 9

Conversion of 18:3n-6 into 20:3n-6 in Leukocytes from Male and Female Volunteers

Incubation time (h)	Male	Female
0	0	0
0.16	24 ± 5	25 ± 4
1	142 ± 60	157 ± 50
24	1479 ± 249	2233 ± 778

Values are expressed in pmol of 20:3n-6 produced/mg cellular protein and represent the mean ± S.D. of 4 volunteers.

REFERENCES

U.S. Pat. No. 3,817,837, June, 1974, Rubinstein et al.
U.S. Pat. No. 3,850,752, Nov., 1974, Schuurs et al.
U.S. Pat. No. 3,939,350, Feb., 1976, Kronick et al.
U.S. Pat. No. 3,996,345, Dec., 1976, Ullman et al.
U.S. Pat. No. 4,275,149, June, 1981, Litman et al.
U.S. Pat. No. 4,277,437, July, 1981, Maggio
U.S. Pat. No. 4,366,241, Dec., 1982, Tom et al.
U.S. Pat. No. 4,399,216, Aug., 1983, Axel et al.
U.S. Pat. No. 4,704,362, Nov., 1987, Itakura et al.
U.S. Pat. No. 4,766,075, Aug., 1988, Goeddal et al.
U.S. Pat. No. 4,784,950, Nov., 1988, Hagen et al.
U.S. Pat. No. 4,801,542, Jan., 1989, Murray et al.
U.S. Pat. No. 4,816,567, Mar., 1989, Cabilly et al.
U.S. Pat. No. 4,935,349, June, 1990, McKnight et al.
U.S. Pat. No. 5,130,238, July, 1992, Malek
International Patent Application No. WO 00/55330, Sept., 2000, Napier J.A.
Patent Cooperation Treaty International Publication No. WO 00/12720, Mar., 2000, Mukerji et al.
European Published Application No. 0320308, June, 1989, Backman et al.
Patent Cooperation Treaty International Publication No. WO 93/05182, Mar., 1993, Bruice
International Patent Application No. WO 88/04300, June, 1988, Cech et al.
Canadian Patent Application No. 2,301,158, Mar., 2000, Winther et al.

Altschul et al., 1990, *J. Molec. Biol.*, 215: 403-410
Ausubel et al., 1994-, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., NY
Been M.D. and Cech T.R., 1986, *Cell*, 47: 207-216
Bennett et al., 1995, *J. Mol. Recognit.*, 8: 52-58
Carillo H. and Lipman D., 1988, *SIAM J. Applied Math.*, 48: 1073
Caskey C.T., 1987, *Science*, 236:1223-1229
Church et al., 1988, *Proc. Natl. Acad. Sci.*, 81: 1991-1995
Cinti et al., 1992, *Prog. Lipid Res.*, 31: 1-51
Connolly B. A., 1987, *Nucl. Acids Res.*, 15: 3131-3139
Copsey et al., 1988, *Genetically Engineered Human Therapeutic Drugs*, Stockton Press, NY
Cotter et al., 1995, *Diabetic Neuropathy: New Concepts and Insights*, Elsevier Science B.V., Amsterdam, pp. 115-120

Cotton et al., 1985, *Proc. Natl. Acad. Sci.*, 85: 4397-4401

Daum et al., 1982, *J. Biol. Chem.*, 257: 13028-13033

Davis et al., 1986, *Basic Methods in Molecular Biology*, Elsevier, NY

Deutcher, M., ed., 1990, *Guide to Protein Purification. Meth. in Enzymol.*, Vol.182

Devereux et al., 1984, *Nucl. Acids Res.*, 12: 387-395

Dines et al., 1993, *Diabetologia*, 36: 1132-1138

Erickson et al., 1992, *Ann. Rep. Med. Chem.*, 27: 271-289

Fields S. and Song O., 1989, *Nature*, 340: 245-246

Flavell et al., 1978, *Cell*, 15: 25-41

Folch et al., 1957, *J. Biol. Chem.*, 226: 497-509

Furuyama K. and Sassa S., 2000, *J. Clin. Invest.*, 105: 757-764

Geever et al., 1981, *Proc. Natl. Acad. Sci.*, 78: 5081-5085

Goding J.W., 1996, *Monoclonal Antibodies: Principles and Practice: Production and Application of Monoclonal Antibodies in Cell Biology, Biochemistry and Immunology, 3rd edition*, Academic Press, NY

Goszcz et al., 1998, *Methods Find. Exp. Clin. Pharmacol.*, 20: 439-445

Gribskov M. and Devereux J., eds., 1991, *Sequence Analysis Primer*, M Stockton Press, NY

Griffin A.M. and Griffin H.G., eds., 1994, *Computer Analysis of Sequence Data, Part I*, Humana Press, NJ

Gyuris et al., 1993, *Cell*, 75: 791-803

Hamy et al., 1997, *Proc. Natl. Acad. Sci.*, 94: 3548-3553

Hanahan et al., 1983, *J. Mol. Biol.*, 166: 557-580

Harlow E. and Lane D., eds., 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor, N.Y.

Haseloff J. and Gerlach W.L., 1988, *Nature*, 334: 585-591

Horrobin D.F. (ed.), 1990, *Omega-6 Essential Fatty Acids: Pathophysiology and Roles in Clinical Medicine*, Wiley-Liss, NY

Houbenweyl et al., 1987, *Methods of Organic Chemistry*, Wansch E. (ed), Vol. 15 I and II, Thieme, Germany

Huse et al., 1989, *Science*, 246: 1275-1281

Hwang et al., 1999, *Proc. Natl. Acad. Sci.*, 96: 12997-13002

Innis M.A. and Gelfand D.H., 1989, *PCR Protocols, A Guide to Methods and Applications*, Innis M.A., Gelfand D.H., Shinsky J.J. and White T.J. (eds), Academic Press, NY, pp. 3-12

Izant J.G. and Weintraub H., 1984, *Cell*, 36: 1007-1015

Jackson et al., 1990, *EMBO J.*, 9: 3153-3162

James et al., 1995, *Plant Cell*, 7: 309-319

Johanson et al., 1995, *J. Biol. Chem.*, 270: 9459-9471

Kawashima et al., 1996, *Biosci. Biotech. Biochem.*, 60: 1672-1676

Kilberg M.S. and Christensen H.N., 1979, *Biochemistry*, 18: 1525-1530

Kohler G. and Milstein C., 1975, *Nature*, 256: 495-497

Kraemer et al., 1993, *J. Lipid Res.*, 34: 663-671

Landegren et al., 1988, *Science*, 241: 1077-1080

Landegren et al., 1989, *Science*, 242: 229-237

Leonard et al., 2000, *Biochem J.*, 350: 765-770

Lesk A.M., ed., 1988, *Computational Molecular Biology*, Oxford University Press, NY

Llewellyn et al., 1987, J. *Mol. Biol.*, 195: 115-123

Lowry et al., 1951, *J. Biol. Chem.*, 193: 265-275

Mack E.W., 1990, *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Easton, Pa., 13th edition

Margolskee et al., 1988, *Mol. Cell. Biol.*, 8: 2837-2847

Mazza G. and Domah B.D. (eds.), 2000, *Herbs, Botanicals, and Teas*, Technomic Publishers, Lancaster, PA

McLaughlin et al., 1988, *J. Virol.*, 62: 1963-1973

Mei et al., 1998, *Biochemistry*, 37: 14204-14212

Merrifield, 1964, *J. Am. Chem. Assoc.*, 85: 2149-2154

Miller et al., 1998, *Nucl. Acids Res.*, 26: 3577-3583

Moss et al., 1987, *Annu. Rev. Immunol.*, 5: 305-324

Myers et al., 1985, *Science*, 230: 1242-1246

Myers et al., 1986, *Cold Spring Harbour Sym. Quant. Biol.*, 51: 275-284

Nilsson T. and Warren G., 1994, *Curr. Opin. Cell Biol.*, 6: 517-521

Oh et al., 1997, *J. Biol. Chem.*, 272: 17376-17384

Okano et al., 1988, *EMBO J.*, 7: 3407-3412

Orkin et al.,1988, *Prog. Med. Genet.*, 7: 130-142

Rasmussen et al., 1987, *Meth. Enzymol.*, 139: 642-654

Riemersma et al., 1986, *Br. Med. J. (Clin. Res. Ed.)*, 292: 1423-1427

Rosenberg et al., 1985, *Nature*, 313: 703-706

Saiki et al., 1985, *Science*, 230: 1350-1353

Saiki, et al., 1986, *Nature*, 324: 163-166

Sambrook et al., 1989, *Molecular Cloning, 2nd Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY

Schatz G. and Kovac L., 1974, *Meth. Enzymol.*, 31A: 627-632

Shanklin et al., 1994, *Biochemistry*, 33: 12787-12794

Singer et al., 1984, *Prostaglandins Leukot. Med.*, 15: 159-165

Smith D.W., ed., 1993, *Biocomputing: Informatics and Genome Project*, Academic Press, NY

Sonnhammer et al., 1998, In *Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology*, AAAI Press, CA, pp. 175-182

Sprecher H., 2000, *Biochim. Biophys. Acta*, 1486: 219-231

Suneja et al., 1990, *Biochim. Biophys. Acta*, 1042: 81-85

Thompson et al., 1994, *Nucl. Acids Res.*, 22: 4673-4680

Toke D.A. and Martin C.E., 1996, *J. Biol. Chem.*, 271: 18413-18422

Tvrdik et al., 2000, *J. Cell. Biol.*, 149: 707-717

Tvrdik et al., 1997, *J. Biol. Chem.*, 272: 31738-31746

Ulmasov et al., 1997, *Science*, 276: 1865-1868

von Heijne G., 1987, *Sequence Analysis in Molecular Biology*, Academic Press, NY

Waldmann T. A., 1991, *Science*, 252: 1657-1661

Wallace et al., 1986, *Cold Spring Harbour Symp. Quant. Biol.*, 51: 257-261

Zaug A.J. and Cech T.R, 1986, *Science*, 231: 470-475

Zaug et al., 1984, *Science*, 224: 574-578

Zaug et al., 1986, *Nature*, 324: 429-433

CLAIMS

We claim:

1. An isolated polynucleotide sequence, comprising a polynucleotide sequence which is selected from the group consisting of:
(a) a sequence comprising SEQ ID NO: 4;
(b) a sequence comprising SEQ ID NO: 8;
(c) a sequence comprising SEQ ID NO: 11;
(d) a sequence which is at least 80% homologous with a sequence of any of (a) to (c);
(e) a sequence which is at least 90% homologous with a sequence of any of (a) to (c);
(f) a sequence which is at least 95% homologous with a sequence of any of (a) to (c);
(g) a sequence which is at least 98% homologous with a sequence of any of (a) to (c);
(h) a sequence which is at least 99% homologous with a sequence of any of (a) to (c);
and;
(i) a sequence which hybridizes to any of (a) to (h) under stringent conditions.

2. An isolated polynucleotide sequence of claim 1, wherein the isolated polynucleotide sequence is cDNA.

3. A vector comprising a polynucleotide sequence of claim 1 in a suitable vector.

4. A host cell comprising a polynucleotide sequence of claim 1 in a host cell which is heterologous to said sequence.

5. An isolated polynucleotide fragment selected from the group consisting of:
(a) a sequence having at least 15 sequential bases of nucleotides of a sequence of claim 1;
(b) a sequence having at least 30 sequential bases of nucleotides of a sequence of claim 1; and
(c) a sequence having at least 50 sequential bases of nucleotides of a sequence of claim 1.

6. A vector comprising a polynucleotide sequence of claim 5 contained in a vector which is heterologous to said sequence.

7. A vector of claims 3 or 6, wherein said vector contains or encodes a tag.

8. An isolated polynucleotide sequence, comprising a polynucleotide sequence which retains substantially the same biological function or activity as or is a functional derivative of a polynucleotide of claim 1.

9. A method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of claim 1, comprising the steps of:
(a) selecting a control animal having said sequence and a test animal having said sequence;
(b) treating said test animal using a compound; and,
(c) determining the relative quantity of an expression product of said sequence, as between said control animal and said test animal.

10. A method of claim 9, wherein said animals are mammals.

11. A method of claim 10, wherein said mammals are rats.

12. A method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of claim 1, comprising the steps of:
(a) selecting a host cell of claim 4;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity of an expression product of said sequence, as between said test group and said control group.

13. A method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of claim 1, comprising the steps of:
(a) selecting a test group having a host cell of claim 4 or a part thereof, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity or relative activity of a product of said sequence or of the said sequence, as between said test group and said control group.

14. An isolated polypeptide comprising an isolated polypeptide selected from the group consisting of:
(a) a sequence comprising SEQ ID NO: 5;
(b) a sequence comprising SEQ ID NO: 9;
(c) a sequence comprising SEQ ID NO: 12;
(d) a sequence which is at least 80% homologous with a sequence of any of (a) to (c);

(e) a sequence which is at least 90% homologous with a sequence of any of (a) to (c);
(f) a sequence which is at least 95% homologous with a sequence of any of (a) to (c);
(g) a sequence which is at least 98% homologous with a sequence of any of (a) to (c); and
(h) a sequence which is at least 99% homologous with a sequence of any of (a) to (c).

15. A host cell comprising a polypeptide sequence of claim 14 in a host cell which is heterologous to said sequence.

16. A process for producing a polypeptide sequence of claim 14 comprising the step of culturing the host cell of claim 15 under conditions sufficient for the production of said polypeptide.

17. An isolated polypeptide sequence, comprising a polypeptide sequence which retains substantially the same biological function or activity as a polypeptide of claim 14.

18. A method for identifying a compound which inhibits or promotes the activity of a polypeptide sequence of claim 14, comprising the steps of:
(a) selecting a control animal having said sequence and a test animal having said sequence;
(b) treating said test animal using a compound;
(c) determining the relative quantity or relative activity of an expression product of said sequence or of the said sequence, as between said control animal and said test animal.

19. A method of claim 18, wherein said animals are mammals.

20. A method of claim 19, wherein said mammals are rats.

21. A method for identifying a compound which inhibits or promotes the activity of a polypeptide sequence of claim 14, comprising the steps of:
(a) selecting a host cell of claim 15;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity or relative activity of an expression product of said sequence or of the said sequence, as between said test group and said control group.

22. A method for identifying a compound which inhibits or promotes the activity of a polypeptide sequence of claim 14, comprising the steps of:

(a) selecting a test group having a host cell of claim 15 or a part thereof, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity or relative activity of a product of said sequence or of the said sequence, as between said test group and said control group.

23. An isolated polynucleotide sequence, comprising a polynucleotide sequence which is selected from the group consisting of:
(a) a sequence comprising SEQ ID NO: 1;
(b) a sequence comprising SEQ ID NO: 2;
(c) a sequence comprising SEQ ID NO: 3;
(d) a sequence comprising SEQ ID NO: 6;
(e) a sequence comprising SEQ ID NO: 7;
(f) a sequence comprising SEQ ID NO: 10;
(g) a sequence comprising SEQ ID NO: 13;
(h) a sequence which is at least 80% homologous with a sequence of any of (a) to (g);
(i) a sequence which is at least 90% homologous with a sequence of any of (a) to (g);
(j) a sequence which is at least 95% homologous with a sequence of any of (a) to (g);
(k) a sequence which is at least 98% homologous with a sequence of any of (a) to (g);
(l) a sequence which is at least 99% homologous with a sequence of any of (a) to (g); and;
(m) a sequence which hybridizes to any of (a) to (l) under stringent conditions.

24. An isolated polynucleotide sequence of claim 23, wherein the isolated polynucleotide sequence is genomic DNA.

25. A vector comprising a polynucleotide sequence of claim 23 in a suitable vector.

26. A host cell comprising a polynucleotide sequence of claim 23 in a host cell which is heterologous to said sequence.

27. A process for producing a polypeptide encoded by a gene operably linked to a polynucleotide sequence of claim 23 comprising the step of culturing the host cell of claim 26 under conditions sufficient for the production of said polypeptide.

28. An isolated polynucleotide fragment selected from the group consisting of:
(a) a sequence having at least 15 sequential bases of nucleotides of a sequence of claim 23;

(b) a sequence having at least 30 sequential bases of nucleotides of a sequence of claim 23; and
(c) a sequence having at least 50 sequential bases of nucleotides of a sequence of claim 23.

29. A vector comprising a polynucleotide sequence of claim 28 contained in a vector which is heterologous to said sequence.

30. An isolated polynucleotide sequence, comprising a polynucleotide sequence which has substantially the same biological function or activity or is a functional derivative of a sequence of claim 23.

31. A method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of claim 23, comprising the steps of:
(a) selecting a control animal having said sequence and a test animal having said sequence;
(b) treating said test animal using a compound; and,
(c) determining the relative quantity of an expression product of an operably linked polynucleotide to said sequence, as between said control animal and said test animal.

32. A method of claim 31, wherein said animals are mammals.

33. A method of claim 32, wherein said mammals are rats.

34. A method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of claim 23, comprising the steps of:
(a) selecting a host cell of claim 26;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity of an expression product of an operably linked polynucleotide to said sequence, as between said test group and said control group.

35. A method for identifying a compound which inhibits or promotes the activity of a polynucleotide sequence of claim 23, comprising the steps of:
(a) selecting a test group having a host cell of claim 26 or a part thereof, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity of an expression product of an operably linked polynucleotide to said sequence, as between said test group and said control group.

36. A composition for treating a PUFA disorder comprising a compound which modulates a sequence according to claims 1, 14 or 23 and a pharmaceutically acceptable carrier.

37. A composition as claimed in claim 36, wherein said disorder is selected from the group consisting eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

38. A composition as claimed in claim 37, wherein said compound is selected from the group consisting of small organic molecules, peptides, polypeptides, antisense molecules, oligonucleotides, polynucleotides, fatty acids and derivatives thereof.

39. The use of a composition as claimed in claim 36 for treating PUFA disorders.

40. The use of claim 41, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

41. A method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting a germline alteration in a sequence of claims 1 or 23 in said subject, comprising comparing the germline sequence of a sequence of claims 1 or 23 from a tissue sample from said subject with the germline sequence of a wild-type of said sequence, wherein an alteration in the germline sequence of said subject indicates the presence of or a predisposition to said PUFA disorder.

42. A method for diagnosing the presence of or a predisposition for a disorder as claimed in claim 41, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis,

endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

43. The method of claims 41 to 42, wherein said comparing is performed by a method selected from the group consisting of immunoblotting, immunocytochemistry, enzyme-linked immunosorbent assay, DNA fingerprinting, in situ hybridization, polymerase chain reaction, reverse transcription polymerase chain reaction, radioimmunoassay, immunoradiometric assay and immunoenzymatic assay.

44. A method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject, comprising comparing the sequence of a polypeptide of claim 14 from a tissue sample from said subject with the sequence of a wild-type of said polypeptide, wherein an alteration in the sequence of said subject as compared to said wild-type indicates the presence of or a predisposition to said PUFA disorder.

45. A method for diagnosing the presence of or a predisposition for a disorder as claimed in claim 44, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

46. The method of claims 44 to 45, wherein said comparing is performed by a method selected from the group consisting of blotting, immunocytochemistry, enzyme-linked immunosorbent assay, DNA fingerprinting, radioimmunoassay, immunoradiometric assay, immunoenzymatic assay and polypeptide microarrays.

47. A method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:

(a) selecting a control animal having said polynucleotide and a test animal having said polynucleotide;

(b) treating said test animal using a compound; and,

(c) determining the relative quantity of an expression product of said polynucleotide, as between said control animal and said test animal.

48. A method of claim 47, wherein said animals are mammals.

49. A method of claim 48, wherein said mammals are rats.

50. A method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a host cell having said polynucleotide, wherein said host cell is heterologous to said polynucleotide;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity of an expression product of said polynucleotide, as between said test group and said control group.

51. A method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a test group having a host cell with said polynucleotide or a portion of said host cell, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity or relative activity of a product of said polynucleotide or of the said polynucleotide, as between said test group and said control group.

52. A method for identifying a compound modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a control animal having said polypeptide and a test animal having said polypeptide;
(b) treating said test animal using a compound;
(c) determining the relative quantity or relative activity of an expression product of said polypeptide or of the said polypeptide, as between said control animal and said test animal.

53. A method of claim 52, wherein said animals are mammals.

54. A method of claim 53, wherein said mammals are rats.

55. A method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a host cell comprising said polypeptide, wherein said host cell is heterologous to said polypeptide;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity or relative activity of an expression product of said polypeptide or of the said polypeptide, as between said test group and said control group.

56. A method for identifying a compound which modulates a PUFA disorder, comprising identifying a compound which modulates the activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a test group having a host cell with said polynucleotide or a portion of said host cell, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity or relative activity of a product of said polypeptide or of the said polypeptide, as between said test group and said control group.

57. A method for identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a control region of a gene selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a control animal having said polynucleotide and a test animal having said polynucleotide;
(b) treating said test animal using a compound; and,
(c) determining the relative quantity of an expression product of an operably linked polynucleotide to said polynucleotide, as between said control animal and said test animal.

58. A method of claim 57, wherein said animals are mammals.

59. A method of claim 58, wherein said mammals are rats.

60. A method for identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a control region of a gene selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a host cell comprising said polynucleotide, wherein said host cell is heterologous to said polynucleotide;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity of an expression product of an operably linked polynucleotide to said polynucleotide, as between said test group and said control group.

61. A method for identifying a compound which modulates the activity of a polynucleotide, wherein the polynucleotide is a control region of a gene selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a test group having a host cell with said polynucleotide or a portion of said host cell, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity of an expression product of an operably linked polynucleotide to said polynucleotide, as between said test group and said control group.

62. A composition for treating a PUFA disorder comprising a compound which modulates a polynucleotide from the coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, and a pharmaceutically acceptable carrier.

63. A composition for treating a PUFA disorder comprising a compound which modulates a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, and a pharmaceutically acceptable carrier.

64. A composition for treating a PUFA disorder comprising a compound which modulates a polynucleotide from the control region selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, and a pharmaceutically acceptable carrier.

65. A composition as claimed in any one of claims 62 to 64, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight

disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

66. A composition as claimed in any one of claims 62 to 64, wherein said compound is selected from the group consisting of small organic molecules, peptides, polypeptides, antisense molecules, oligonucleotides, polynucleotides, fatty acids and derivatives thereof.

67. The use of a composition as claimed in any one of claims 62 to 64 for treating PUFA disorders.

68. The use of claim 67, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

69. A method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting a germline alteration in a polynucleotide representing the coding sequence selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, from said subject, comprising comparing the germline sequence of said polynucleotide from a tissue sample from said subject with the germline sequence of a wild-type of said polynucleotide, wherein an alteration in the germline sequence of said subject indicates the presence of or a predisposition to said PUFA disorder.

70. A method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting a germline alteration in a polynucleotide representing the control region selected from the group consisting of ELG1, ELG2, ELG3 and ELG5 in said subject, comprising comparing the germline sequence of said polynucleotide from a tissue sample from said subject with the germline sequence of a wild-type of said polynucleotide, wherein an alteration in the germline sequence of said subject indicates the presence of or a predisposition to said PUFA disorder.

71. A method for diagnosing the presence of or a predisposition for a disorder as claimed in any one of claims 69 to 70, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

72. The method of claims 69 to 71, wherein said comparing is performed by a method selected from the group consisting of immunoblotting, immunocytochemistry, enzyme-linked immunosorbent assay, DNA fingerprinting, in situ hybridization, polymerase chain reaction, reverse transcription polymerase chain reaction, radioimmunoassay, immunoradiometric assay and immunoenzymatic assay.

73. A method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject, comprising comparing the sequence of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, and ELG7, from said subject with the sequence of a wild-type of said polypeptide, wherein an alteration in the sequence of said subject as compared to said wild-type indicates the presence of or a predisposition to said PUFA disorder.

74. A method for diagnosing the presence of or a predisposition for a disorder as claimed in claim 73, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

75. The method of claims 73 to 74, wherein said comparing is performed by a method selected from the group consisting of immunoblotting, immunocytochemistry, enzyme-linked immunosorbent assay, DNA fingerprinting, radioimmunoassay, immunoradiometric assay, immunoenzymatic assay and polypeptide microarrays.

76. A method for identifying a compound which inhibits or promotes the overall activity of two or more polynucleotides, wherein the polynucleotides are control regions of two or

more different genes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a host cell having said polynucleotides, wherein said host cell is heterologous to said polynucleotides;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantities of expression products of operably linked polynucleotides to said polynucleotides, as between said test group and said control group.

77. A method for identifying a compound which inhibits or promotes the overall activity of two or more polynucleotides, wherein the polynucleotides are from control regions of said polynucleotides, selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of:
(a) selecting a test group having a host cell with said polynucleotide or a portion of said host cell, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantities of expression products of operably linked polynucleotides to said polynucleotides, as between said test group and said control group.

78. A method for identifying a compound which inhibits or promotes the activity of two or more polynucleotides, wherein the polynucleotides are coding sequences selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence, comprising the steps of:
(a) selecting a host cell having said polynucleotides, wherein said host cell are heterologous to said polynucleotides;
(b) cloning said host cell and separating said clones into a test group and a control group;
(c) treating said test group using a compound; and
(d) determining the relative quantity or relative activity of an expression product of said polynucleotide, as between said test group and said control group.

79. A method for identifying a compound which inhibits or promotes the activity of two or more polynucleotides, wherein the polynucleotides are coding sequences selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence, comprising the steps of:

(a) selecting a test group having a host cell with said polynucleotide or a portion of said host cell, and selecting a suitable control group;
(b) treating said test group using a compound; and
(c) determining the relative quantity or relative activity of an expression product of said polynucleotide, as between said test group and said control group.

80. A method for identifying a compound which inhibits or promotes the activity of a mammalian delta-5-desaturase enzyme and one or more enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, within the same host system, comprising the steps of:
(a) providing a host system containing nucleic acid sequences which encode for a mammalian delta-5-desaturase and one or more mammalian elongase enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence;
(b) contacting the host system with a test component;
(c) simultaneously evaluating the enzymatic activities of the delta-5-desaturase and the elongase enzymes, wherein a measurable difference in a level of lipid metabolites or associated cofactors in the presence of the test component compared to a control under identical conditions but in the absence of the test component is an indicator of the ability of the test component to modulate delta-5-desaturase and/or elongase enzyme activity; and
(d) identifying as said compound a test component which exhibits said ability.

81. A method for identifying a compound which inhibits or promotes the activity of a mammalian delta-6-desaturase enzyme and one or more enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, within the same host system, comprising the steps of:
(a) providing a host system containing nucleic acid sequences which encode for a mammalian delta-6-desaturase and one or more mammalian elongase enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence;
(b) contacting the host system with a test component;
(c) simultaneously evaluating the enzymatic activities of the delta-6-desaturase and the elongase enzymes, wherein a measurable difference in a level of lipid metabolites or associated cofactors in the presence of the test component compared to a control under

identical conditions but in the absence of the test component is an indicator of the ability of the test component to modulate delta-6-desaturase and/or elongase enzyme activity; and
(d) identifying as said compound a test component which exhibits said ability.

82. A method for identifying a compound which inhibits or promotes the activity of a mammalian delta-5- and delta-6-desaturase enzyme and/or one or more enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, within the same host system, comprising the steps of:
(a) providing a host system containing nucleic acid sequences which encode simultaneously for a mammalian delta-5-desaturase, a mammalian delta-6-desaturase and one or more mammalian elongase enzymes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 or ELG7, operably associated with promoter regions, wherein the promoter regions are effective to initiate, terminate or regulate the level of expression of the nucleic acid sequence;
(b) contacting the host system with a test component;
(c) simultaneously evaluating the enzymatic activities of the delta-5-desaturase, the delta-6-desaturase and the elongase enzymes, wherein a measurable difference in a level of lipid metabolites or associated cofactors in the presence of the test component compared to a control under identical conditions but in the absence of the test component is an indicator of the ability of the test component to modulate delta-5- and/or delta-6-desaturase and/or elongase enzyme activity; and
(d) identifying as said compound a test component which exhibits said ability.

83. A composition for treating a PUFA disorder comprising a compound which modulates two or more human polynucleotides from control regions selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, ELG7, delta-5-desaturase, delta-6-desaturase and a pharmaceutically acceptable carrier.

84. A composition as claimed in claim 83, wherein said compound is selected from the group consisting of small organic molecules, peptides, polypeptides, antisense molecules, oligonucleotides, polynucleotides, fatty acids and derivatives thereof.

85. The use of a composition as claimed in claim 84 for treating PUFA disorders.

86. The use of claim 85, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic

fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

87. A method for detecting the presence of or the predisposition for a PUFA disorder, said method comprising determining the level of expression of two or more expression products of genes selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6, ELG7, delta-5-desaturase, delta-6-desaturase, in a subject relative to a predetermined control level of expression, wherein any modified expression of said expression products as compared to said control is indicative of the presence of or the predisposition for a PUFA disorder.

88. A method of claim 87, wherein said disorder is selected from a group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

89. A method of claims 87 to 88, wherein said method is selected from the group consisting of immunoblotting, immunocytochemistry, enzyme-linked immunosorbent assay, in situ hybridization, reverse transcription polymerase chain reaction, radioimmunoassay, immunoradiometric assay, immunoenzymatic assay and polynucleotide and polypeptide microarrays.

90. An antibody immunoreactive with a polypeptide of claim 14 or an immunogenic portion thereof.

91. An antibody immunoreactive with an elongase polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, or an immunogenic portion thereof.

92. A method for screening a medium for an elongase polypeptide of claim 14 or selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising;

(a) labelling an antibody of any one of claims 90 to 91 with a marker molecule to form a conjugate;
(b) exposing said conjugate to said medium; and
(c) determining whether there is binding between said conjugate and a biomolecule in said medium, wherein said binding indicates the presence of said polypeptide.

93. A method for screening a medium for an elongase polypeptide of claim 14 or selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising:
(a) exposing an antibody of claims 91 to 92 to said medium;
(b) exposing said antibody to a marker molecule; and
(c) determining whether there is binding between said marker molecule and a biomolecule in said medium, wherein said binding indicates the presence of said polypeptide.

94. A composition as claimed in claim 36, wherein said compound is selected from the group in claim 90.

95. A composition as claimed in any one of claims 62 to 64, wherein said compound is selected from the group consisting of antibodies against ELG1, ELG2, ELG3 and ELG5.

96. The use of a composition as claimed in any one of claims 94 to 95 for treating a PUFA disorder.

97. The use of claim 96, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

98. A compound identified by the methods of any one of claims 9 to 13, 18 to 22, 31 to 35, 47 to 61 or 76 to 82.

99. The use of a compound as claimed in claim 98 for treating a PUFA disorder.

100. The use as claimed in claim 99, wherein said disorder is selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

101. A method for diagnosing the presence of or a predisposition for a PUFA disorder in a subject by detecting alterations as compared to wild-type in the elongation of PUFA in a peripheral blood leukocyte obtained from said subject.

102. A method for monitoring the development of a PUFA disorder in a subject by detecting alterations as compared to previous samples in the elongation of PUFA in a peripheral blood leukocyte obtained from said subjects.

103. A method for assessing the efficacy of test compounds on a PUFA disorder in a subject by assessing alterations as compared to previous samples in the elongation of PUFA in a peripheral blood leukocyte obtained from said subject.

104. The use of pebulate sulphoxide for the treatment of a disease selected from the group consisting of eczema, cardiovascular, inflammation, Sjögren's syndrome, gastrointestinal disorders, viral diseases and postviral fatigue, body weight disorders, psychiatric disorders, cancer, cystic fibrosis, endometriosis, pre-menstrual syndrome, alcoholism, congenital liver disease, Alzheimer's syndrome, hypercholesterolemia, autoimmune disorders, atopic disorders, acute respiratory distress syndrome, articular cartilage degradation, diabetes and diabetic complications.

105. A method for identifying a compound which modulates a biological activity of a polypeptide selected from the group consisting of ELG1, ELG2, ELG3, ELG4, ELG5, ELG6 and ELG7, comprising the steps of: (a) providing an assay which measures a biological activity of the selected polypeptide; (b) treating the assay with a compound; and (c) identifying a change in the biological activity of the selected polypeptide, wherein a difference between the treated assay and a control assay identifies the compound as modulator of the polypeptide.

106. The method of claim 105, wherein the selected polypeptide is provided in an assay format selected from among a purified protein, reconstituted protein, cell extract and a whole cell assay.

107. The composition as claimed in any one of claims 37 and 65, wherein the cardiovascular disorder is selected from the group consisting of hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease, cerebrovascular disease and peripheral vascular disease.

108. The use as claimed in any one of claims 40, 68, 86, 97, 100, and 104, wherein the cardiovascular disorder is selected from the group consisting of hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease, cerebrovascular disease and peripheral vascular disease.

109. The method as claimed in any one of claims 42, 45, 71, 74, and 88, wherein the cardiovascular disorder is selected from the group consisting of hypertriglyceridemia, dyslipidemia, atherosclerosis, coronary artery disease, cerebrovascular disease and peripheral vascular disease.

110. The composition as claimed in any one of claims 37 and 65, wherein the inflammation is selected from the group consisting of sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne.

111. The use as claimed in any one of claims 40, 68, 86, 97, 100, and 104, wherein the inflammation is selected from the group consisting of sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne.

112. The method as claimed in any one of claims 42, 45, 71, 74, and 88, wherein the inflammation is selected from the group consisting of sinusitis, asthma, pancreatitis, osteoarthritis, rheumatoid arthritis and acne.

113. The composition as claimed in any one of claims 37 and 65, wherein the body weight disorder is selected from the group consisting of obesity, cachexia and anorexia.

114. The use as claimed in any one of claims 40, 68, 86, 97, 100, and 104, wherein the body weight disorder is selected from the group consisting of obesity, cachexia and anorexia.

115. The method as claimed in any one of claims 42, 45, 71, 74, and 88, wherein the body weight disorder is selected from the group consisting of obesity, cachexia and anorexia.

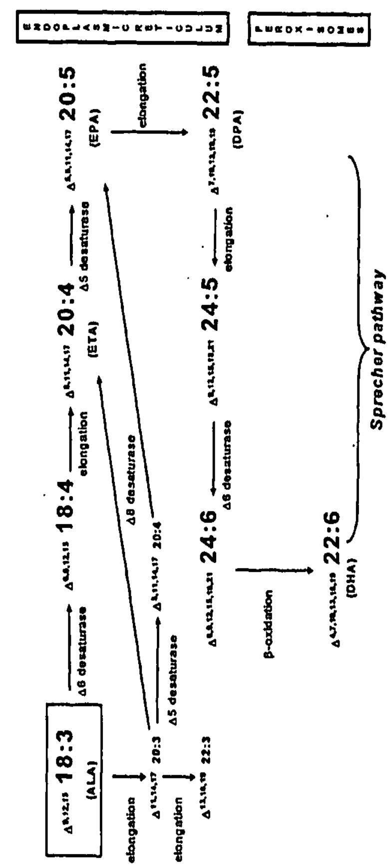


FIGURE 1

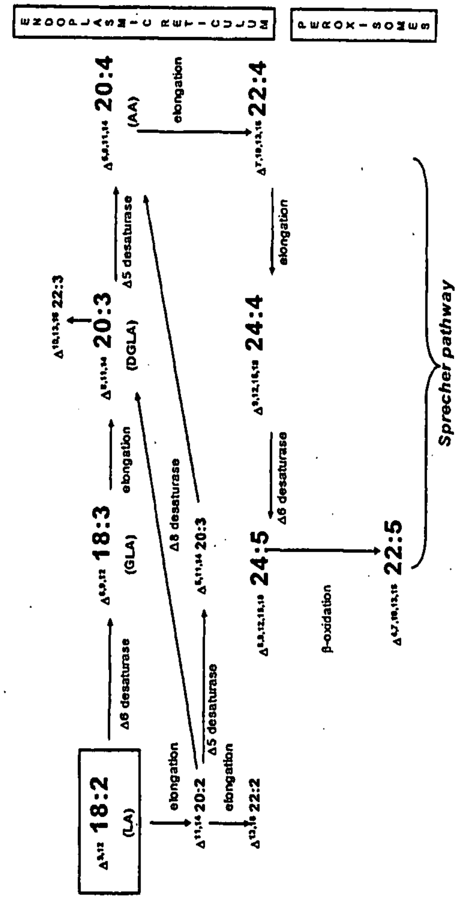


FIGURE 2

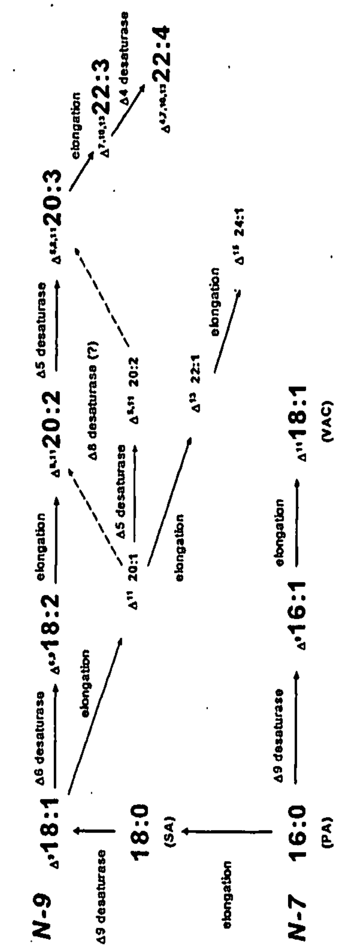


FIGURE 3

FIGURE 4

```
ELG1 ------------------MEAVVNLYQEVMKHADPRIQGYPLMGSPLLMTSILLT  37
ELG2 ---------------MEHFDASLSTYFKALLGPRDTRVKGWFLLDN-YIPTFICSV  40
ELG3 ------------MEHLKAFDDEINAFLDNMFGPRDSRVRGWFMLDS-YLPTFFLTV  43
ELG4 -------------MAFSDLTSRTVHLYDNWIKDADPRVEDWLLMSSPLPQTILLGF  43
ELG5 -------------------MNMSVLTLQEYEFEKQFNENEAIQWMQENWKKSFLFSA  38
ELG6 ---------------MVTAMNVSHEVNQLFQPYNFELSKDMRPFFEEYWATSFPIAL  42
ELG7 MGLLDSEPGSVLNVVSTALNDTVEFYRWTWSIADKRVENWPLMQS-PWPTLSIST  54

            *    *
ELG1 YVYFVLSLGPRIMANRKPFQLRGFMIVYNFSLVALSLY-IVYEFLMSGWLSTYTW  91
ELG2 IYLLIVWLGPKYMRNKQPFSCRGILVVYNLGLTLLSLY-MFCELVTGVWEGKYNF  94
ELG3 MYLLSIWLGNKYMKNRPALSLRGILTLYNLGITLLSAY-MLAELILSTWEGGYNL  97
ELG4 YVYFVTSLGPKLMENRKPFELKKAMITYNFFIVLFSVY-MCYEFVMSGWGIGYSF  97
ELG5 LYAAFIFGGRHLMNKRAKFELRKPLVLWSLTLAVFSIFGALRTGAYMVYILMTKG  93
ELG6 IYLVLIAVGQNYMKERKGFNLQGPLILWSFCLAIFSILGAVRMWGIMGTVLLTGG  97
ELG7 LYLLFVWLGPKWMKDREPFQMRLVLIIYNFGMVLLNLF-IFRELFMGSYNAGYSY 108

                                       **   *   **    ***         *  **
ELG1 RCDPVDYSNSPEALRMVRVAWLFLFSKFIELMDTVIFILRKKDGQVTFLHVFHHS 146
ELG2 FCQGTRTAG-ESDMKIIRVLWWYYFSKLIEFMDTFFFILRKNNHQITVLHVYHHA 148
ELG3 QCQDLTSAG-EADIRVAKVLWWYYFSKSVEFLDTIFFVLRKKTSQITFLHVYHHA 151
ELG4 RCDIVDYSRSPTALRMARTCWLYYFSKFIELLDTIFFVLRKKNSQVTFLHVFHHT 152
ELG5 LKQSVCDQGFYNGPVSKFWAYAFVLSKAPELGDTIFIILRKQ--KLIFLHWYHHI 146
ELG6 LKQTVCFINFIDNSTVKFWSWVFLLSKVIELGDTAFIILRKR--PLIFIHWYHHS 150
ELG7 ICQSVDYSNNVHEVRIAAALWWYFVSKGVEYLDTVFFILRKKNNQVSFLHVYHHC 163
```

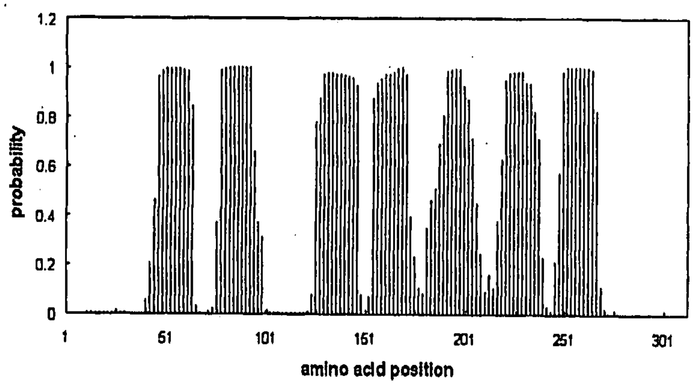
FIGURE 4 (continued)

```
                          *       *      *     **   **   *
ELG1 VLPWSWWWGVKIAPGGMGSFHAMINSSVHVIMYLYYGLSAFGPVAQPYLWWKKHM 201
ELG2 SMLNIWWFVMNWVPCGHSYFGATLNSFIHVLMYSYYGLSSV-PSMRPYLWWKKYI 202
ELG3 SMFNIWWCVLNWIPCGQSFFGPTLNSFIHILMYSYYGLSVF-PSMHKYLWWKKYL 205
ELG4 IMPWTWWFGVKFAAGGLGTFHALLNTAVHVVMYSYYGLSALGPAYQKYLWWKKYL 207
ELG5 TVLLYSWYSYKDMVAGGGWF-MTMNYGVHAVMYSYYALRAA--GFRVSRKFAMFI 198
ELG6 TVLVYTSFGYKNKVPAGGWF-VTMNFGVHAIMYTYYTLKAA--NVKPPKMLPMLI 202
ELG7 TMFTLWWIGIKWVAGGQAFFGAQLNSFIHVIMYSYYGLTAFGPWIQKYLWWKRYL 218

      *   *   *                                       **   *
ELG1 TAIQLIQFVLVSLHISQYYFMSSCNYQYPVIIHLIWM--YGTIFFMLFSNFWYHS 254
ELG2 TQGQLLQFVLTIIQTSCGVIWP-CTFPLGWLYFQIG---YMISLIALFTNFYIQT 253
ELG3 TQAQLVQFVLTITHTMSAVVKP-CGFPFGCLIFQSS---YMLTLVILFLNFYVQT 256
ELG4 TSLQLVQFVIVAIHISQFFFMEDCKYQFPVFACIIMS--YSFMFLLLFLHFWYRA 260
ELG5 TLSQITQMLMGCVVNYLVFCWMQHDQCHSHFQNIFWSSLMYLSYLVLFCHFFFEA 253
ELG6 TSLQILQMFVGAIVSILTYIWRQDQGCHTTMEHLFWSFILYMTYFILFAHFFCQT 257
ELG7 TMLQLIQFHVTIGHTALSLYTD-CPFPKWMHWALIA---YAISFIFLFLNFYIRT 269

      *
ELG1 YTKG-KRLPRALQQNGAPGIAKVKAN----------------------- 279
ELG2 YNKKGASRRKDHLKDHQNGSMAAVNGHTNSFSPLENNVKPRKLRKD--- 299
ELG3 YRKK--PMKKDMQEPP--AGKEVKNGFSKAYFTAANGVMNKKAQ---- 296
ELG4 YTKG-QRLPKTVKN----GTCKNKDN----------------------- 281
ELG5 YIG---KMRKTTKAE---------------------------------- 265
ELG6 YIRP--KVKAKTKSQ---------------------------------- 270
ELG7 YKEP---KKPKAGKTAMNGISANGVSKSEKQLMIENGKKQKNGKAKGD 314
```

FIGURE 5



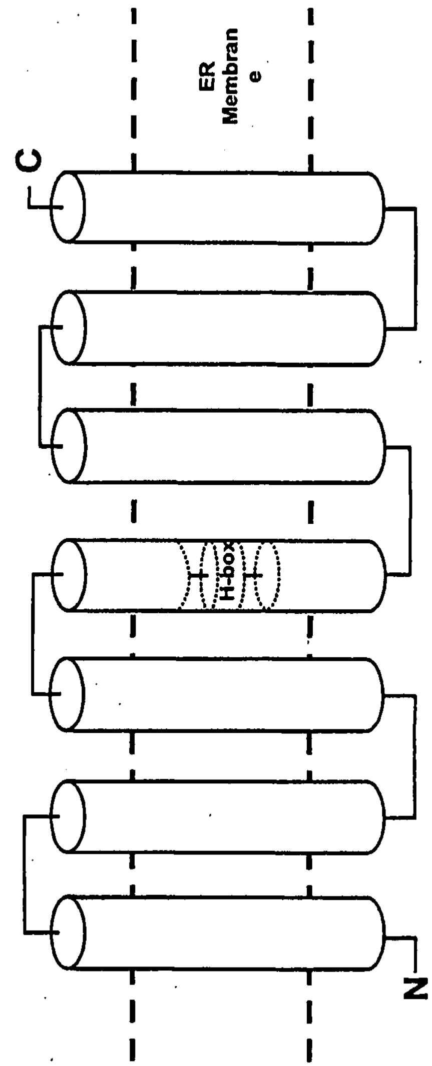


FIGURE 6

FIGURE 7

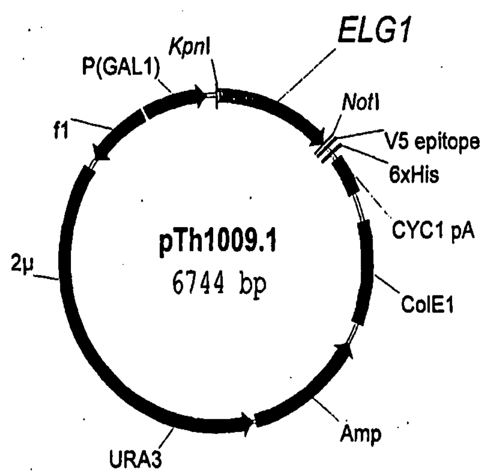


FIGURE 8

1	TACAGGCTCG	TGAGGCTTCC	CTCCCGCTAA	GACCAGTGCG	CCCTCAGCAC
51	ACGCAGTGTG	GTCTCGCCCG	CCGCTCTGCG	CTCGCCCTGC	AGGAGAGGGA
101	GCTCTTTGAA	GGCAAGGCCG	AACCTCCCCC	GAGCCCTGAG	CTGGGCCTGC
151	CGCCACAGAT	GTGCAGTCCT	GCCGGGGAGC	AGTCACCCGG	GGACAGGGCC
201	GGGCCCCGGG	CTGCACGTCG	GGAAGAGACA	GCGTGCTCCT	GAGGTGGCCA
251	GGCCGCTGCA	ACTGGCCAGG	GCGGGCCCGG	GCGGCGAGGG	AAGGGGTGGG
301	AAGCCCGGGC	CGCGGCGCTT	CCTGCTGGGA	CCCGGCGGCA	CGCCCCTGCC
351	CCCGCCCCGG	CCGAGCCTGC	GCTGCCGGCC	TCCGGCCCTG	CCGGCCGCCC
401	AATCAGCGGG	CGCCCCCCGC	GCGGCCCGCC	CCTCCCCCTC	TGGTGACAGA
451	AAGTCGGCCC	AGCAGATGAG	GAAGTGGCAG	GCAGGCAGGC	TGGCCCCGGG
501	GACTTCTCTC	TGGCCCTGCT	CCCTCCGAGC	GCTCCGCCGT	TGCCCGCCTG
551	GCCCCTACGG	GTGAGTCTGG	ACCTTCCACG	GACTCTCCAC	GTGCCGGCGC
601	CCCCTGCCTG	GCCAGCCCGG	CCCAGCCCGG	CCCAGCCCTG	CCCTGCCCTG
651	CCCAGGCTGT	GGGCGAGGGT	GTTCCCGGGG	CCAGTGGGTG	GGAGGTCCCA
701	GCTCCCTGGG	GCCGGGCCTC	GCCAGCACCC	TCCCTCCCCC	ACACCCCCGT
751	CTCTGGCCCC	CATTTGCCTA	CACCCGGGCC	TTCCTCCACC	ACCCCTGCAT
801	TTACCTCTCT	CCCTCCTCCT	CTCCCCTCCC	TCCCCCCGCT	ACCCTAACTT
851	TGCCAGGCAC	CTTTTCCCTT	CCATCCATCT	TAAAGGAAGG	AAGGGACGGG
901	CTGAGTTCCC	CGACGAGAGA	CACACCCAGA	TTTTCCTGCA	GCTTGGGGAG
951	AGGTCCTCCC	AGGAGCCTTG	GTCCCTCCTG	GCCTGCCGG	

FIGURE 9

```
  1 CGAGGGTGGG CTTCTGCCAC CCAAATGCGG CCACAGACTC CTGCCACGCC
 51 TGGCAGTAAA AAAACCAGAG TTCAGGGCAT CGACAACTTC ACCGGGGCTA
101 TTGCGCAGGC TCTGCGTTCC ACGCAGGCTT ATTAGGAAGA AAGGGGAAAA
151 AAATTTCCCA GAGACACGTG GAACCGAGGG GCCAACCCCG GCCTAGGCTC
201 TCCACCGCAT CGGATTCTGG AATTTACGAT CACGAAAGTT CTATTGTCCC
251 GCGATTGGCT CCCGGGCCGC ATGACATCAT AGCGCTTGAT TCATCCTTCG
301 GGTCCCGATT GGCTGGCCGC GCCATTGTGA CGTCACGGTC AGCCCACGTT
351 CTGATTGTAG ATAGCCGGCG CCTTCCTCTT CCCATCGCGC GGGTCCTAGC
401 CACCGGTGTC TCCTTCTACA TCCGCCTCTG CGCCGGCTGC CACCCGCGCT
451 CCCTCCGCCG CCGCCGCCTT GCTGCTGCTC AAAGCTGCTG CCGCCCCTTG
501 GGCTAAAAG
```


FIGURE 10

```
   1 CCGGTACCTA CAGTTACTCA CTCTGCTACT GCACAAAACT CTGCAAGGGC
  51 TCCCACACCG CCCCAGGTGT GGGATGCTAA GTGTATGGTG CAGGTACCTC
 101 CGTGCACAGC CACACGGGCT GCTCTCAACC CCAATAAACA TGTTTACCAC
 151 ATGAGCCTCA CATGTGGTAA ACATTTTTTT TTTTTTTTTT TTTTTTTTGA
 201 GACAGGGTCT CATTCTGTCG CCCAGGCTGG AGTGCAGTGG CGTGATCTCG
 251 GCTCACTGCA GCCTCCACCT CCAGGGCTCA AGCCATCCTT CACCTCAGCC
 301 TCCCGAGTGG CTGGGATCAC AGGCGCAGGC CACCACACCC AGCTAATTTT
 351 TGTATTTTTT ATTTAAGAGG CGGGGTTTCG CCATGTTGCC CAGGCTGGTC
 401 CCGAACTCCT GACCTCAAGT GATTCGCCTG CCTCAGCCTC CCAAAGTGCT
 451 GGGATTACAG GAGGGAACCA CCACGCCCGC CAACTTCCCA TGCTTGAGGG
 501 AGAAATGGAA GAAAGTTCAT GTAATACTCA GGCAAGTCCA ATTTTTTCGA
 551 CGTCTTTCAC TTGGGCCACA CACACAACTA AAGTAACTAG AAGCGCAGGC
 601 TCTAGGAGGC CACCGTTCTG TTCACAGTGA AGAGGGTGCG CTCACCGTTG
 651 GTCGTGTCCG CTGGAAGCCC CGCGTCAGGC CGGGAGCGGG ACAGAGACTC
 701 TTGCTCAGGG CCGTTATCCG AACTGATCCG CTTCCCACCG CACCCCCAGA
 751 GAAACCCACC CAACCCCCTA AACCTAAGAA ACCCAGACTG CGCAAACCTG
 801 CAGGAACAGA GCCATTTCCC CCTAATGTGT GCTTCAAACC CACCGAAGCC
 851 CAACTGTAAG CAAGACCAGC GTGCCCGCCC TGCACGATAC TGCTTCTCCC
 901 CGCAGCAGCG GCTGCCGATC TGGGCAGCGG GTGGGTATTC CTGGGGCTCC
 951 GTGGACGTTG AGCCGCCGCG CGAAACCGGC GCCGGCTGGA CCTGCAAATC
1001 GCCGCCCGGC CGGCAGGGGA CGCCGCGGAC GCGAGGGCGA GGTCGGTCGC
```

FIGURE 10 (continued)

```
1051  CCAGGAGGGG GCGCGCGAGG CCGCAGGGGC GGGGGGCGCC GCCTCACTTG
1101  CCCTGCGCCC CTCCCCCGCG CGCCCTCCTG GCGCGGCGGC CGGCGAGGCC
1151  CCTGTGGGAG AGGGGGCGGG GACGAAACGG CCCCGAGGCT CGGAGCGCCG
1201  CGCGGCGGCG GCGCGAGCCC GAGGGGGCGG GGAGGCGCGG GCGGGTGTGC
1251  GCGCGCCGGG CGTGGGTGTG GGTGGGGGTA ACCGGCGCGG GCGCCGAGAT
1301  AGCGCCGGGC AGAGGGACCC GGCTACCCTG GACAGCGCAT CGCC
```


FIGURE 11

```
                                       30
5'  ATG GCC TTC AGT GAT CTT ACA TCG AGG ACT GTG CAT CTT TAT GAT AAT TGG ATC
    M   A   F   S   D   L   T   S   R   T   V   H   L   Y   D   N   W   I

        60                                         90
    AAA GAT GCT GAT CCA AGA GTT GAA GAT TGG CTC CTC ATG TCC TCG CCT CTG CCA
    K   D   A   D   P   R   V   E   D   W   L   L   M   S   S   P   L   P

                120                                    150
    CAA ACC ATC CTC CTA GGA TTC TAT GTC TAT TTT GTC ACT TCC TTG GGA CCA AAG
    Q   T   I   L   L   G   F   Y   V   Y   F   V   T   S   L   G   P   K

                            180                                210
    CTC ATG GAA AAT CGC AAG CCC TTT GAA CTC AAG AAA GCA ATG ATA ACG TAC AAT
    L   M   E   N   R   K   P   F   E   L   K   K   A   M   I   T   Y   N

                                        240                                270
    TTT TTC ATA GTA CTC TTT TCT GTG TAT ATG TGT TAT GAG TTT GTG ATG TCT GGC
    F   F   I   V   L   F   S   V   Y   M   C   Y   E   F   V   M   S   G

                                       300
    TGG GGT ATA GGT TAT TCA TTT CGA TGT GAC ATT GTT GAC TAT TCA CGG TCA CCC
    W   G   I   G   Y   S   F   R   C   D   I   V   D   Y   S   R   S   P

        330                                        360
    ACA GCT TTG AGG ATG GCA CGT ACC TGC TGG CTT TAT TAC TTC TCC AAA TTT ATT
    T   A   L   R   M   A   R   T   C   W   L   Y   Y   F   S   K   F   I

                390                                        420
    GAG CTA TTA GAT ACG ATC TTT TTT GTT CTG CGC AAG AAA AAT AGC CAA GTG ACT
    E   L   L   D   T   I   F   F   V   L   R   K   K   N   S   Q   V   T

                            450                                480
    TTC CTT CAT GTA TTC CAT CAT ACC ATC ATG CCG TGG ACC TGG TGG TTT GGA GTC
    F   L   H   V   F   H   H   T   I   M   P   W   T   W   W   F   G   V

                                        510                                540
    AAA TTT GCT GCA GGT GGT TTG GGA ACA TTC CAT GCC CTT CTA AAT ACA GCT GTA
    K   F   A   A   G   G   L   G   T   F   H   A   L   L   N   T   A   V

                                            570
    CAT GTA GTC ATG TAT TCC TAC TAT GGA CTT TCT GCA TTG GGG CCA GCC TAC CAG
    H   V   V   M   Y   S   Y   Y   G   L   S   A   L   G   P   A   Y   Q
```

FIGURE 11 (continued)

```
    600                                         630
AAG TAT TTG TGG TGG AAA AAA TAT TTG ACA TCA TTA CAG CTT GTC CAG TTT GTT
K   Y   L   W   W   K   K   Y   L   T   S   L   Q   L   V   Q   F   V

        660                                 690
ATT GTC GCC ATC CAC ATA AGC CAG TTC TTT TTC ATG GAG GAT TGC AAG TAT CAG
I   V   A   I   H   I   S   Q   F   F   F   M   E   D   C   K   Y   Q

                720                                 750
TTT CCA GTC TTT GCG TGC ATC ATT ATG AGT TAC AGT TTC ATG TTT CTG CTG CTC
F   P   V   F   A   C   I   I   M   S   Y   S   F   M   F   L   L   L

                        780                                         810
TTT CTC CAT TTT TGG TAC CGT GCT TAC ACC AAA GGT CAG AGG TTG CCC AAA ACT
F   L   H   F   W   Y   R   A   Y   T   K   G   Q   R   L   P   K   T

                                840
GTG AAA AAT GGA ACT TGC AAA AAC AAA GAT AAT TGA 3'
V   K   N   G   T   C   K   N   K   D   N   *
```


FIGURE 12

```
   1 TGCGCCTGGC TGAACACTAC ATTTTTTTTA CTTCTTTATT CATGTATTGT
  51 CTGTCATCTC CAACTAGAAT GAACGTATAG TCCCTGAGAA CGGGGAATTT
 101 GTTATCTATT GAAACTTCAG GGCCTGGAAC ATAGCAGCAC TCCAGTATTT
 151 GTTAAATAAA TGAATCCATT TGAGCTTCTG CATATTTGAA ATTTCATAAG
 201 TATATATAAA TGGTAAATTG TGATAGACTC AAAGGCTAGT ATCATTAGGC
 251 AATTGTCTCC CGTTCCCAAA AGACTTCCTA AGTCTACTAA ATGATCTGTT
 301 TTTAATATGA AAGCAAAGTT ATCTAAAAGA AAGGAGAAAT CTTTAGTTTT
 351 TTTGACTTCG AGATTCTTTG CAATTTAAGC TTTTTTTTTT TTTTTTTTTT
 401 TTTTTTTTTG CTTTTCTTTC AATGGACACT TTCGAAGTTT TACATAAAAA
 451 CATTAAAACC TCTTGTTTAA TGTAGTGGGA TTAAGCTGCC GAAGGCAATC
 501 CCTACATGTG AGGAAAATAT GCTTCCGACA CCCCAATTTT TTTTTTCTCC
 551 CTACCCATCC TCTCTGGTGG TCCTGACGCT CCCAGCCCCT TTTTGTGTTT
 601 CTTGATTCCA TGCTGAGAAC TCGCAATACA AACTCAAAGC CCACATTTGT
 651 GAGGTGGTTG GGTCAGGACT GCAACTAAAA ATGATTATTG TTTTTTAGGT
 701 TTCTGGACAG TTCAACACCA GCCTTTGGTT TTGCCTCAGA AGCAGGGAAC
 751 TTCTCTAGGC CCCTATTTTG CCTTTCAGCT ATTGATGATC CAAATCATAC
 801 CAGCGATTAG GAGGATCATT ACCAGACACA AGGCCAGGTA CGTTTAAAAA
 851 ATAAATAAAC CAAGCGCAGG TGCACACTCC GAACGCTCAT CCCCACCCCC
 901 ACTTTCCAAT CCAACAGTAG GTAACGAGAA ATGAATTTTC TAGACTTTTT
 951 TTCCTGCAGC AGTTGCTGTT ACCAGAAACA AAGTTAGATG ATATACAATC
1001 TAATCTTCAT TGCTCTAAAA GTCCTCTCCC CATGCCCCCC AGGCTGCCTC
```

FIGURE 12 (continued)

```
1051  AATTCTCTAG TTTCTTATTC CTTATAAGCA GGGGATGGAG CTGAACCAAG
1101  TCGGCCTTCC CCTCCCAGGG CCTTCTCCTC TTGGTCTGGC TTCCATTTCA
1151  GATGCGAATT AACCCTCCCA ATACCCTTTC AGAAGCAAGG AGTCCCCTTT
1201  TTCTCCGCCT CCAGCCTCAG CTAGGTTTTC CTCATTTCGG ATTTTTCTAC
1251  AGCTCATTCC CAAATGAGTC ACGCATGACG ACAATTTCCA CTCTGCTATG
1301  TCAGCCTGGA GATGTCCCCC AAGTGATGGC ATCTGCTCTC GGAAAGAAAG
1351  GTCATCGGTG CCACGACCAG CCCCGCTAAC CCAGAGCGGC CGGTGGGCCC
1401  CAGTCCCGAG AGTCAGGGCG CGCGGCGGAG GCGAGGCCGG GGCGGCCTCC
1451  GCCCTCCCGG CCGCTCCCCC TCGCGCCGCC CCGGCTCCTC CCTCCGGCCC
1501  TCGGCGGGCA CCTGGCGGCG GCGGGCAGGG GGCGGCGCTG CGCGCGTCAC
1551  GCGGCTGGGT GGGATAGCGG GCAGGTGACA CCCGGCGGCC TCCTCCCCTT
1601  TCCAACCCAG TCGGCGGCCG GGACAGCAGG GGCCGCTGTG AGGAGCTCCG
1651  CGCTCGCGCT GCCAGTCGCC GCCCTCTCTC CCGCGCGCGC CCGGCGCTTC
1701  GGCTCCGCTC CCTGTGCGGT GAGTGCGGGG TTCCAGGCCG GCGGGCAGGG
1751  GCCAAACTTT CCCGGCGCGC GGAGGAGAAG AGACTGGGGA GGGAGGCAGA
1801  GCCGAGGGGA ACGGCGTCGG GAGTGGCCGG ATGGAGGAAC TTGGGCGCGG
1851  CGCGCGAGAA GTGGGACCCG GGTGCGGGGG CCCCGGGAGC GGGGCCAGGC
1901  CCTCCCTGGG CTCGGGAGGC GCTTGGGAAG TTCTGTCCCC GCTGCCTGCG
1951  CGTGGGGAGG ACCGAGGCCC TTTTCGCCGG AGCGCGGGGC CGCGGCGCTC
2001  ACCTGCGCCT TCTCGGGAGC CCCCACCCGG CAGCATCCCG AAGGGAAGGT
2051  CGGGCCCGGT GGGCGCGCTG CGGAGCGGAG CCTGGACTGG GGTCCCGCGC
```


FIGURE 12 (continued)

```
2101  GGCGCTGGCC CTGCGGAGCG GAGCGGGAGG GGCAGAGGTG CTCGCCGGCG
2151  GGACTGGGAG GGAGAAGGAC CTGCTCGACC TTGGACGCGG AGGTCATTTT
2201  CCCAGCTCCG GGGTCTGGCC TCGCTAGCCA CCCCCCCAAA TTCCGGAGCC
2251  CCTTTCTTTC TGTTTCCTTC CTTCCCCTTT GGGCGCTTTT TTTGCTCCCG
2301  CGGCCAGATG AACTTGGGGC GCTGTCCCTT CGGCTCCCCG AGCCGCATCC
2351  TGTCTTGGTG GCTGCTGCTG GCCGGGAGGA GGCTGATGAA TACAGAGCCG
2401  TGGAACAGGT CGTGCCGGAG ATGGAAACAG GAAAGCCTGT TGTTTTGTCG
2451  TCCCAG
```

FIGURE 13

```
   1 GTGAGCCACC ACCGCGGCCG GTCCCTTCCT CCTTTAAAAA TTTTTCTCCC
  51 AGTTCCCACT TTTTGTGGGT TAGAGGCATC TAAATTGAAT GAAAGTACCC
 101 TTTTTGGACT ACTGGGGAGG TGGGGGGATG TTCTCAGAAG GGGAATTTTC
 151 TTTCTGGTCC TAATATCCAC CTAATTTTTA AAAGCAGGGC TCCTTATTAT
 201 TTTGTAAAGT TTACAATTAC ATCATTAGAT ACTTCCATGT CTCATATTTC
 251 ATTTTTCCAA ACTCTTGGGG GAAATGAGTG GAGGGATGGA TGGAATAGAA
 301 AATAGTTTTT CCTCTTGGAG GCTGAGGGCC CAGTAGGGGT CAACAGTACA
 351 TTCAGCCCTC TCCTCACATA TTCTGTTCTA CCTACAAGTA CAGCAAGTAA
 401 AGCCAAATTT CTCATGCATG CAAATAAAGT TTTTGCATTT GGCCAGTCGG
 451 TCCAGTTCTC CTGTCAGCTT CCTTCCCCAC TCTGCCTCTG TTCATTAATC
 501 CCCCCCTTCC CGGTACCTAA ACCCTCCACC TAACCCAGCC CTTTCTTCCA
 551 CTTCCGGCTA CTAGCCTCTC TCGCCTATCC ACTATCCTCA CACTCAGCAT
 601 CCCCTGTCTG TACGAGATTA AGGAGCTCTG CCGTCCGCAG GGCCTGGGTT
 651 AGCGTGAATC TAAGCCAGAG CTCCCGGGTG GGGGTGGGGG TAGGGGTGGG
 701 GGTGGTCCCA GAGGTAGGGC GAGGAGGTGG GAAGCGTATT CCCTTCACTG
 751 GTGATCTCAA CGTAGATTTG CCCGGAGTTC TCTTGCAAGA GAGCTGGCAG
 801 GTTTTACTAT TTCCCAATCG TTTACTCGCC AAGCTCTCGG GTCCACGCGC
 851 CGCGGGGATG CGCCCTGCAA GCTGAAACTT CATTCAAAGC AAGGCGGCCC
 901 ACGAGGTTGG GCTTAGGGGA TCTGGATGAC CTCCAGGCCA CTTCCTTTCT
 951 CTCTGCGCCC TTCCCCCACT CTTCCAACCA CCTTCGCTGT AAACAAAACT
1001 GTCCCCCCCG GGCGGAGAGA GGTCGCGCTC TTTCGCACAC TCCCTCGCCA
```


FIGURE 13 (continued)

```
1051 AGGGTTAATT TCTCAAATCG CACGAGGGGG AGGAGATTTC CCTGTAGACG
1101 AGTAAAAAGG GTGATGGACA AACGTGCGGG CACTAAGACC GCAAGGCATT
1151 CATTTCCTCC TACGGTGGAT GCGGACGCCG GGAGGAGGAG AGCCCCAGAG
1201 AGAGGAGCTG GGAGCGGAGG CGCAGGCAAT GCTCAGCCCT GGATGTAGCT
1251 GAGAGGCTGG GAGAAGAGAC GACCGCTGGA GACCGAGCGG CGTGGGGAAG
1301 ACCTAGGGGG GTGGGTGGGG GAAGCAGACA GGAGAACACT CGAAATCAAG
1351 CGCTTTACAG ATTATTTTAT TTTGTATAGA GAACACGTAG CGACTCCGAA
1401 GATCAGCCCC A
```

FIGURE 14

```
                                              30
5'  ATG GTC ACA GCC ATG AAT GTC TCA CAT GAA GTA AAT CAG CTG TTC CAG CCC TAT
    M   V   T   A   M   N   V   S   H   E   V   N   Q   L   F   Q   P   Y

        60                                          90
    AAC TTC GAG CTG TCC AAG GAC ATG AGG CCC TTT TTC GAG GAG TAT TGG GCA ACC
    N   F   E   L   S   K   D   M   R   P   F   F   E   E   Y   W   A   T

            120                                         150
    TCA TTC CCC ATA GCC CTG ATC TAC CTG GTT CTC ATC GCT GTG GGG CAG AAC TAC
    S   F   P   I   A   L   I   Y   L   V   L   I   A   V   G   Q   N   Y

                        180                                         210
    ATG AAG GAA CGC AAG GGC TTC AAC CTG CAA GGG CCT CTC ATC CTC TGG TCC TTC
    M   K   E   R   K   G   F   N   L   Q   G   P   L   I   L   W   S   F

                                    240                                 270
    TGC CTT GCA ATC TTC AGT ATC CTG GGG GCA GTG AGG ATG TGG GCC ATT ATG GGG
    C   L   A   I   F   S   I   L   G   A   V   R   M   W   G   I   M   G

                                                300
    ACT GTG CTA CTT ACC GGG GGC CTA AAG CAA ACC GTG TGC TTC ATC AAC TTC ATC
    T   V   L   L   T   G   G   L   K   Q   T   V   C   F   I   N   F   I

        330                                                 360
    GAT AAT TCC ACA GTC AAA TTC TGG TCC TGG GTC TTT CTT CTC AGC AAG GTC ATA
    D   N   S   T   V   K   F   W   S   W   V   F   L   L   S   K   V   I

            390                                                 420
    GAA CTC GGA GAC ACA GCC TTC ATC ATC CTG CGT AAG CGG CCA CTC ATC TTT ATT
    E   L   G   D   T   A   F   I   I   L   R   K   R   P   L   I   F   I

                        450                                         480
    CAC TGG TAC CAC CAC AGC ACA GTG CTC GTG TAC ACA AGC TTT GGA TAC AAG AAC
    H   W   Y   H   H   S   T   V   L   V   Y   T   S   F   G   Y   K   N

                                    510                                 540
    AAA GTG CCT GCA GGA GGC TGG TTC GTC ACC ATG AAC TTT GGT GTT CAT GCC ATC
    K   V   P   A   G   G   W   F   V   T   M   N   F   G   V   H   A   I

                                                570
    ATG TAC ACC TAC TAC ACT CTG AAG GCT GCC AAC GTG AAG CCC CCC AAG ATG CTG
    M   Y   T   Y   Y   T   L   K   A   A   N   V   K   P   P   K   M   L
```


FIGURE 14 (continued)

```
    600                                             630
CCC ATG CTC ATC ACC AGC CTG CAG ATC TTG CAG ATG TTT GTA GGA GCC ATC GTC
P   M   L   I   T   S   L   Q   I   L   Q   M   F   V   G   A   I   V

            660                                             690
AGC ATC CTC ACG TAC ATC TGG AGG CAG GAT CAG GGA TGC CAC ACC ACG ATG GAA
S   I   L   T   Y   I   W   R   Q   D   Q   G   C   H   T   T   M   E

                        720                                         750
CAC TTA TTC TGG TCC TTC ATC TTG TAT ATG ACC TAT TTC ATC CTC TTT GCC CAC
H   L   F   W   S   F   I   L   Y   M   T   Y   F   I   L   F   A   H

                                    780                                 810
TTC TTC TGC CAG ACC TAC ATC AGG CCC AAG GTC AAA GCC AAG ACC AAG AGC CAG
F   F   C   Q   T   Y   I   R   P   K   V   K   A   K   T   K   S   Q

TGA  3'
*
```

FIGURE 15

```
   1 GATTAGCTGT CAGGCTATAT ATGGAGCCAT CAGGAACCAC TGAAGGTTTT
  51 TTTTTTTTTT TTTTTTTTTG AGACGGAGTC TCACTCTGTC ACCCAGGCTG
 101 GAGTGCAGTG GCACGATCTC TGCTCACTGC AAGCTCTGCC TCCCAGGTTC
 151 ACGCCATTCT CCTGCCTCAG CCTCCCGAGT AGCTGGGACT ACAGGCGCCT
 201 GCCACCACGC CCGGCTAATT TTTTGTATTT TTTAGTAGAG ACGGGGTTTG
 251 ACGGTGTTAG CCAGGATGGT CTCGATCTCC TGACCTCATG ATCTGCCCGC
 301 CTCGGCCTCC CAAGGTGCTG GGATTACAGG CGTGAACCAC CGTGCCCGGC
 351 CGAACCACTG AAGGTTTTTA AGCAGGAAAG CAGAGCTGTT TTCTGGATGA
 401 GCAAACAGAA AGTAGTGGTT TTCCAAGTAC AGTCTGAGAC AACCTATAGG
 451 ACCAGAATCT CTGCAGTTGA GGCTCAGGAA TCTGGTAATC AGCCAGGTAT
 501 AGGAACTCTT TTCTGATTGC AATGCAGTGA AGAGCAGAAG CACTGTATTA
 551 GAGAAAGAGG CAGTGCAACC AGGTAACGTG ACCAGGTGAG AAGTGATGAG
 601 GTACAGAGAC AAAGAGATGC ACTTTTGAGT CACTTAGATG GCACTGATAG
 651 GACTTCCACT ACACCCTCGC ATAGACAGTG GCTGAGGTTC AGGAAATAGA
 701 GCTGGGGTTC CTACTTGGAT CCTCTGGCTC TAGAGCTTTA CTGCACATAG
 751 CCATTTATAC CCACATCTTG ATTTTAATTA TTTTATATCT ATGTTTCTTA
 801 GCACTTTTTG CAAATTTCCA CCTTATCTCA AACTGCCCTC AAGCCTTGTA
 851 TTTCTCCTTC GCTTTCATAA AACCTAGGAA AGAAATAAGG GACAGCCAAG
 901 TAAAACTTTT AAAAGTTTTA GAACATTTAT TTCTTTGGGG CTGGTTACAC
 951 AGGCGAGAAA GAAGTAGATT TGGTTAGGGA GAGAAAACAA CAGGCCTTGG
1001 GGAGATACAC TGGCTCTCCC CCTCCCTAAA CCCTAAGAGG CCTCCAGGAA
```


FIGURE 15 (continued)

```
1051 ACCTGAAGAC AATAATTCCA GAAGCCCAGA GGGTGACCCC ATTTCCTCTC
1101 TCCATGGTTA TTACTGTCAG TCTGGAGCAG TTCAGGAATT CAGGAAACTA
1151 TAAAGAAACC ACAACAGCCT CAACAACCCA AACATCAACA TCAACAACCT
1201 CAACAATAAA ACTCCTTAAA ATTCATCTCC TTCCACCCAC TCACAACCGC
1251 AGACTCGAAG CTAGGAGGTG GAAGGGACTA CAGAAGCTCT GCGTTGCCCA
1301 GGTTAGTATT TGCTCATCAC AGGCCTGGGT TTCCCAGGAT CTCAGGGAGC
1351 CTGGAAACTG ACGCCTCCAT TTCTGGGTGG GAGCACCAAA GCCTAAGGAC
1401 ACCTTTCCTC TCTCTTCACT GCTAAGCAGG TCAAGATTAA AGCAAACCGA
1451 GGCAAAGGCC ACGGTTGACA GTTCCAAGGG AACCCGCAAG GCCGCACAGG
1501 ATGGGGTGGA CGTTTTACGG GAGAAAAGGC TGGGGAAGTG GGCGGGCGAT
1551 GGCCTACGAC GGGACTTGGG GCGGGGTGTG CGAAACGCCT GGCAGGCGGG
1601 CCCTTGAGTA TGACCAATCA GAATGCGGAC TGCGTCCCAG GGGCGGAGCA
1651 GAGGCGTATC TTGGTCGAGA TTGGATAGCG GCGGGGCGCA GGAAAGAGGT
1701 CGCGCCAGCC CGGGCAGGCA GCTTTGCAAG TCCGCGTTAT ATATCGCAGT
1751 GGCTGCGCCC GGGATAGCTG GCTGCGCCGC CGCGCACATG CCTAGGTTCG
1801 ACGCCCTCCT CCCTTTGCCC AGGAGTTCCT TCTGTCCCGG CTCTGTTCCG
1851 TCTCGCCCCG AGGTTCACGC CATCCTCGGA GCCCCAGCCT TTCACCCAGC
1901 GCCTCCAAGC TTTGGACCTT GACTTCTGCA AAACTAG
```

FIGURE 16

```
                                          30
5'  ATG GGG CTC CTG GAC TCG GAG CCG GGT AGT GTC CTA AAC GTA GTG TCC ACG GCA
    M   G   L   L   D   S   E   P   G   S   V   L   N   V   V   S   T   A

        60                                      90
    CTC AAC GAC ACG GTA GAG TTC TAC CGC TGG ACC TGG TCC ATC GCA GAT AAG CGT
    L   N   D   T   V   E   F   Y   R   W   T   W   S   I   A   D   K   R

                120                                     150
    GTG GAA AAT TGG CCT CTG ATG CAG TCT CCT TGG CCT ACA CTA AGT ATA AGC ACT
    V   E   N   W   P   L   M   Q   S   P   W   P   T   L   S   I   S   T

                            180                                     210
    CTT TAT CTC CTG TTT GTG TGG CTG GGT CCA AAA TGG ATG AAG GAC CGA GAA CCT
    L   Y   L   L   F   V   W   L   G   P   K   W   M   K   D   R   E   P

                                        240                                 270
    TTT CAG ATG CGT CTA GTG CTC ATT ATC TAT AAT TTT GGG ATG GTT TTG CTT AAC
    F   Q   M   R   L   V   L   I   I   Y   N   F   G   M   V   L   L   N

                                                    300
    CTC TTT ATC TTC AGA GAG TTA TTC ATG GGA TCA TAT AAT GCG GGA TAT AGC TAT
    L   F   I   F   R   E   L   F   M   G   S   Y   N   A   G   Y   S   Y

        330                                             360
    ATT TGC CAG AGT GTG GAT TAT TCT AAT AAT GTT CAT GAA GTC AGG ATA GCT GCT
    I   C   Q   S   V   D   Y   S   N   N   V   H   E   V   R   I   A   A

                390                                             420
    GCT CTG TGG TGG TAC TTT GTA TCT AAA GGA GTT GAG TAT TTG GAC ACA GTG TTT
    A   L   W   W   Y   F   V   S   K   G   V   E   Y   L   D   T   V   F

                            450                                     480
    TTT ATT CTG AGA AAG AAA AAC AAC CAA GTT TCT TTC CTT CAT GTG TAT CAT CAC
    F   I   L   R   K   K   N   N   Q   V   S   F   L   H   V   Y   H   H

                                        510                                 540
    TGT ACG ATG TTT ACC TTG TGG TGG ATT GGA ATT AAG TGG GTT GCA GGA GGA CAA
    C   T   M   F   T   L   W   W   I   G   I   K   W   V   A   G   G   Q

                                                    570
    GCA TTT TTT GGA GCC CAG TTG AAT TCC TTT ATC CAT GTG ATT ATG TAC TCA TAC
    A   F   F   G   A   Q   L   N   S   F   I   H   V   I   M   Y   S   Y
```


FIGURE 16 (continued)

```
    600                                             630
TAT GGG TTA ACT GCA TTT GGC CCA TGG ATT CAG AAA TAT CTT TGG TGG AAA CGA
Y   G   L   T   A   F   G   P   W   I   Q   K   Y   L   W   W   K   R

            660                                             690
TAC CTG ACT ATG TTG CAA CTG ATT CAA TTC CAT GTG ACC ATT GGG CAC ACG GCA
Y   L   T   M   L   Q   L   I   Q   F   H   V   T   I   G   H   T   A

                        720                                             750
CTG TCT CTT TAC ACT GAC TGC CCC TTC CCC AAA TGG ATG CAC TGG GCT CTA ATT
L   S   L   Y   T   D   C   P   F   P   K   W   M   H   W   A   L   I

                                    780                                 810
GCC TAT GCA ATC AGC TTC ATA TTT CTC TTT CTT AAC TTC TAC ATT CGG ACA TAC
A   Y   A   I   S   F   I   F   L   F   L   N   F   Y   I   R   T   Y

                                                840
AAA GAG CCT AAG AAA CCA AAA GCT GGA AAA ACA GCC ATG AAT GGT ATT TCA GCA
K   E   P   K   K   P   K   A   G   K   T   A   M   N   G   I   S   A

    870                                             900
AAT GGT GTG AGC AAA TCA GAA AAA CAA CTC ATG ATA GAA AAT GGA AAA AAG CAG
N   G   V   S   K   S   E   K   Q   L   M   I   E   N   G   K   K   Q

            930
AAA AAT GGA AAA GCA AAA GGA GAT TAA  3'
K   N   G   K   A   K   G   D   *
```

FIGURE 17

1 GGAAATACCT GAAGCTGTTT TAACAATTTC TCCTTGTATT AAGTATTATG

51 CTGCAGTTTT GCGTGTGTGA ATGGAAGTAT GGGTAGAGAT CTGTTCTCCC

101 TAAAAACTCC AGGATTCCAC AATATAGAAA TAGTAATCAA ATTTTTAGGT

151 GAAGCTCGAA CTAATCCGAA CTTTGTTAGA TCATCACTGT AAATGAATGG

201 GTATTTATCC ACTCCCTAAA TGAAGAGACT TGACTGGATT TCTTTTTTTT

251 ATATAGCTAC TAGAATCTGT TACACATAAT TTAGGATTGA GACTTGAGAA

301 ATTGTCATTC CAATCCAGAA AACTTTAGAT TTGCAAATAT ATTTGACAAA

351 TTAATAAATT AACATTTTAT TTGGTTAATT TCAAGAATAG GGCATTTAAA

401 GAAGTCTGTG TTTGCTTTAG TTCGGCAATA AAGTTCCTGC CACTCACAAT

451 AATCCTTATT ATTCTCTGAA AGACATGTTA TATTTTTGTC ATCATAAATA

501 TTTATTAATT ACTGTTTATA GCACTGGGTT AGGTACTCAT CAAGCAACCA

551 AAAATAATTC TTACCATCTA GGATGCTTCC AATATAAAAT ATAGACAATA

601 TATAACCAGG TCAATTGGGA AATAGATCAT TTCAGTATGA TAAAAGATAG

651 TATTCACATT AACAGTGTGA AAGGGCAGGA ACAATAAGAC ACTTGACTCA

701 CTGGTCTTTA AAATGTAGCA TCCAAAATGA GCAAGTGGAG AAAAGGTTAA

751 ACAAGTAGGT GACACATTTA AAAAACAAGT AGATGAAAGG ACTATTCTCA

801 AAAATCTTGT TTTATGTGAG AAACCATCAA ATTATGAATT CCAAGTACTG

851 TATTTTTTTT ACTTTTCAAG GGTAGGCTCT CCTATACCTT ATCTAAACAA

901 TTTTTCAAAA TAGCCACAAT TACTTTGTTT TCCTCTCTAC ACTAAATTGC

951 CCTTTGCCTC TTGAGCGATT ATCTTTTTCA GATTCACCTC AACTTCTTCA

1001 GGTTCAAGCG GACTTCACCT GTAAGCCCCT CTCGGTTCTC CCTCTTCTCT

FIGURE 17 (continued)

1051 GAACTACTAA TGGCCTAATT TAGCACAATT ATATTGCTTT GTTCATTCCA

1101 TGTATAGTAA AAGAGTCTAC AAAACACATG CAAGCATTCA TGCAATTATA

1151 TGTTGATTTG TTCATGGGTC GACCCCAAAG TCTATTCTCC ATCGCTGAAG

1201 CATGGAAGAC AAATACCCTT CACTTCTTCA GAGGCATAAC ACATGCACTT

1251 CTCTTGTCAT GGTGACAGGC ATGTGCTGGT GGAGGTCAAA GAAACAGGAA

1301 CACAAGTGAA ATCGAGGTGA GTGTCAGGTA AGGACCAAAG CACCACGCCT

1351 ACCTCATCTT TGCCCACAGA ACACCCATTC TTCCCGTGTC CTGTTTCCCA

1401 GGACGTATCC GGGGCGGATA AGAAATCACC CGTGGGGAGG CGGTGAACTC

1451 CTCCGCAGGG GCCGATGCCC GGGACAGGGG CGGGGAAGGC TAATGAGGCG

1501 ACTTGTGCGG GGAGGGGCCA AGGAGGAGCC CAGGTGTCCC GCTCCCGCTC

1551 GACGGCGCGC GCCTGCGCGA GCCCAGTTGG CGTCGCACCC TTGAGCGCAG

1601 CATCCCTACG CCAGCGAGTC CCAATACTAG GGAGGGAGGG AGGGAGGAGG

1651 GGCGGCCGGC CCCCCGCCCC CGCGCGCGGC CACGTGACGC CGGCTGAGGA

1701 GATTGGAGGG GCGGCTGCGC GAGGCTGCAG ACTGGTGCAG CGCACTGTGC

1751 TGGCGGCTGG GCCTCCTCCA CCTCCTCGTC TTTCTCCCGG GAACCTTGAC

1801 GACGCCTTCC GCTTGGCCCT GCCTTCTGCC GCATCCCCGC CGCCGCGGCG

1851 CCTTGAGGAG CAGGAGAAGA CGCAGCCGGG CCGCCGCCGT TAGAGGGGTT

1901 CCCGGCCGCC GCTCGCCCCG TCGGCCGCCA CCGCCTCCGG GGTCAGCCCT

1951 CTCTCTGGGT CTCCGCTTTC TCCTGCCGCC AGCGCCCGCT CATCGCCGCG

FIGURE 18

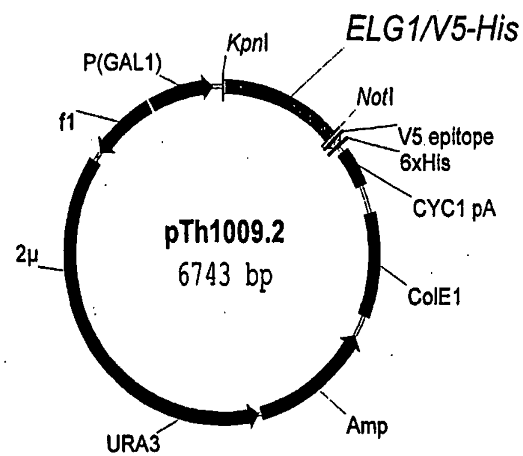


FIGURE 19

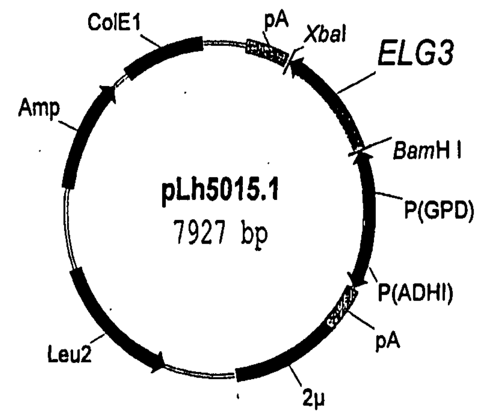


FIGURE 20

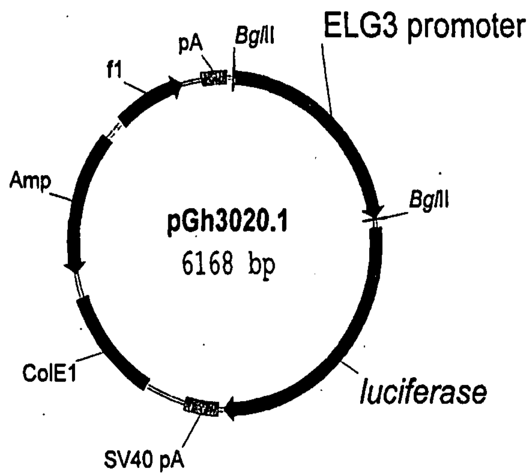


FIGURE 21

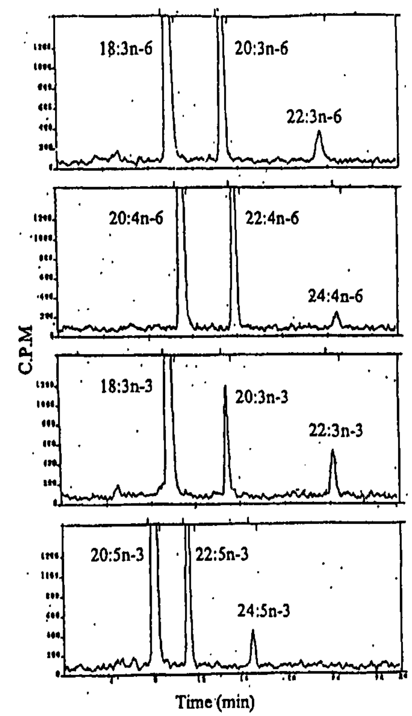


FIGURE 22

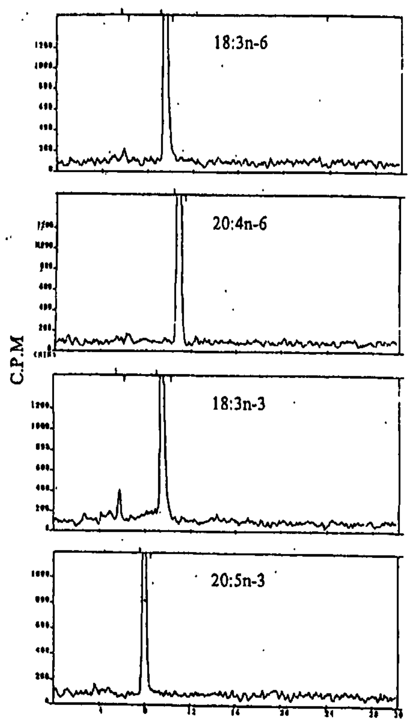


FIGURE 23

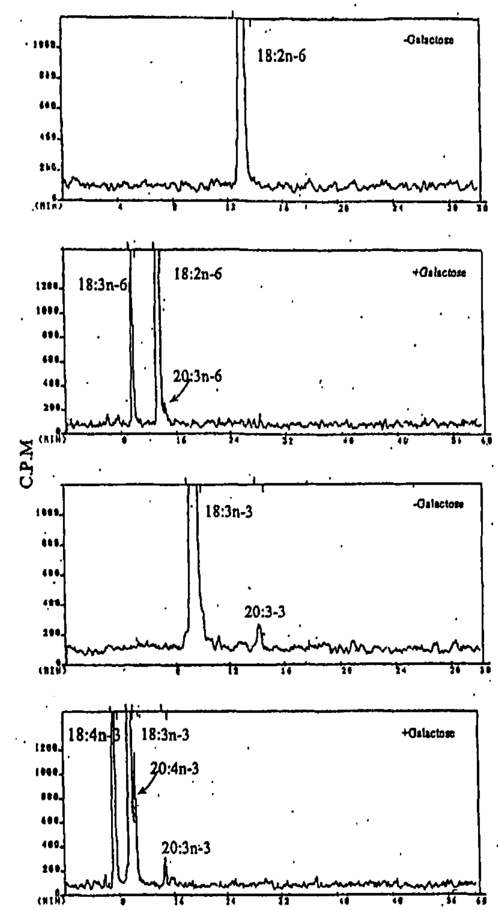


FIGURE 24

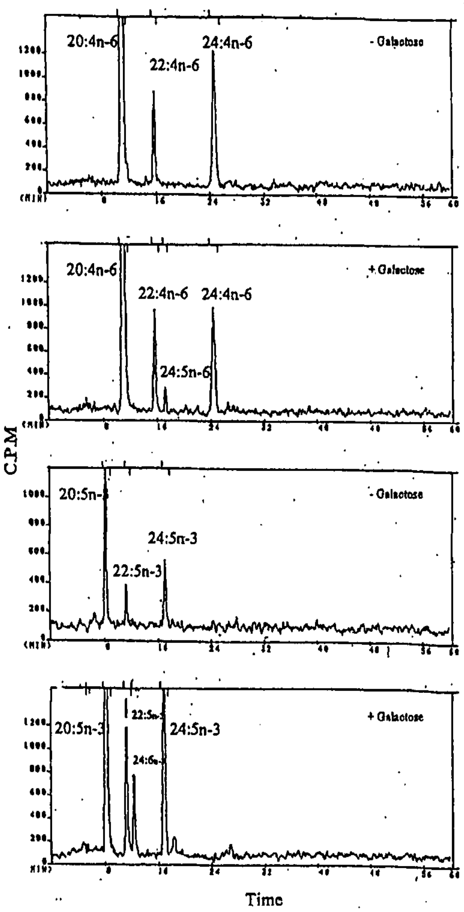


FIGURE 25

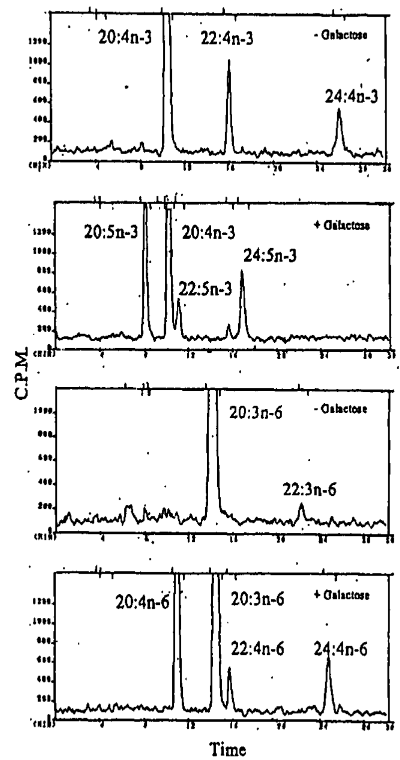
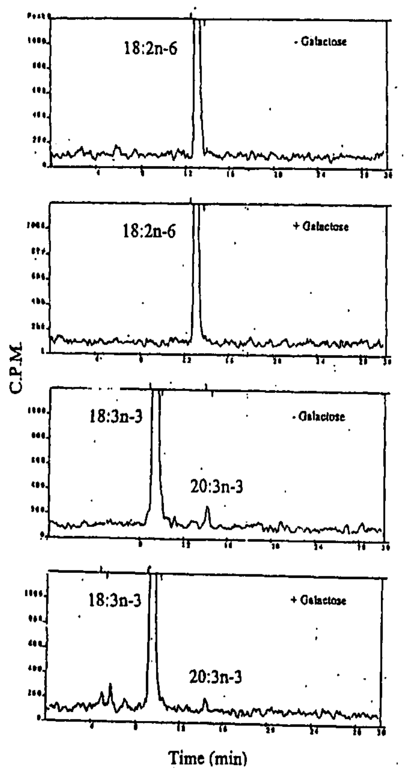


FIGURE 26



SEQUENCE LISTING

<110> Winther, Michael D
Knickle, Leah C
Haardt, Martin
Allen, Stephen J
Ponton, Andre
De Antueno, Roberto J
Jenkins, D K
Nwaka, Solomon O
Goldberg, Y Paul

<120> Human Elongase Genes, Uses Thereof, and Compounds for
Modulating Same

<130> 42320-0008

<140> filed herewith
<141> 2001-11-29

<150> US 60/253,728
<151> 2000-11-29

<160> 58

<170> PatentIn Ver. 2.1

<210> 1
<211> 989
<212> DNA
<213> human

<400> 1
tacaggctcg tgaggcttcc ctcccgctaa gaccagtgcg ccctcagcac acgcagtgtg 60
gtctcgcccg ccgctctgcg ctcgccctgc aggagaggga gctctttgaa ggcaaggccg 120
aacctccccc gagccctgag ctgggcctgc cgccacagat gtgcagtcct gccggggagc 180
agtcacccgg ggacagggcc gggccccggg ctgcacgtcg ggaagagaca gcgtgctcct 240
gaggtggcca ggccgctgca actggccagg gcgggcccgg gcggcgaggg aaggggtggg 300
aagcccgggc cgcggcgctt cctgctggga cccggcggca cgcccctgcc cccgccccgg 360
ccgagcctgc gctgccggcc tccggccctg ccggccgccc aatcagcggg cgcccccggc 420
gcggcccgcc cctcccctc tggtgacaga aagtcggccc agcagatgag gaagtggcag 480
gcaggcaggc tggccccggg gacttctctc tggccctgct ccctccgagc gctccgccgt 540
tgcccgcctg gcccctacgg gtgagtctgg accttccacg gactctccac gtgccggcgc 600
cccctgcctg gccagcccgg cccagcccgg cccagccctg ccctgccctg cccaggctgt 660
gggcgagggt gttcccgggg ccagtgggtg ggaggtccca gctccctggg gccgggcctc 720
gccagcaccc tccctccccc acacccccgt ctctggcccc catttgccta cacccgggcc 780
ttcctccacc acccctgcat ttacctctct ccctcctcct ctcccctccc tccccccgct 840
accctaactt tgccaggcac cttttccctt ccatccatct taaaggaagg aagggacggg 900
ctgagttccc cgacgagaga cacacccaga ttttcctgca gcttggggag aggtcctccc 960
aggagccttg gtccctcctg gcctgccgg 989

<210> 2
<211> 509
<212> DNA
<213> human

<400> 2
cgagggtggg cttctgccac ccaaatgcgg ccacagactc ctgccacgcc tggcagtaaa 60
aaaaccagag ttcagggcat cgacaacttc accggggcta ttgcgcaggc tctgcgttcc 120

```
acgcaggctt attaggaaga aaggggaaaa aaatttccca gagacacgtg gaaccgaggg 180
gccaaccccg gcctaggctc tccaccgcat cggattctgg aatttacgat cacgaaagtt 240
ctattgtccc gcgattggct cccgggccgc atgacatcat agcgcttgat tcatccttcg 300
ggtcccgatt ggctggccgc gccattgtga cgtcacggtc agcccacgtt ctgattgtag 360
atagccggcg ccttcctctt cccatcgcgc gggtcctagc caccggtgtc tccttctaca 420
tccgcctctg cgccggctgc cacccgcgct ccctccgccg ccgccgcctt gctgctgctc 480
aaagctgctg ccgccccttg ggctaaaag                                   509

<210> 3
<211> 1344
<212> DNA
<213> human

<400> 3
ccggtaccta cagttactca ctctgctact gcacaaaact ctgcaagggc tcccacaccg 60
ccccaggtgt gggatgctaa gtgtatggtg caggtacctc cgtgcacagc cacacgggct 120
gctctcaacc ccaataaaca tgtttaccac atgagcctca catgtggtaa acattttttt 180
tttttttttt tttttttga gacagggtct cattctgtcg cccaggctgg agtgcagtgg 240
cgtgatctcg gctcactgca gcctccacct ccagggctca agccatcctt cacctcagcc 300
tcccgagtgg ctgggatcac aggcgcaggc caccacaccc agctaatttt tgtatttttt 360
atttaagagg cggggtttcg ccatgttgcc caggctggtc ccgaactcct gacctcaagt 420
gattcgcctg cctcagcctc ccaaagtgct gggattacag gagggaacca ccacgcccgc 480
caacttccca tgcttgaggg agaaatggaa gaaagttcat gtaatactca ggcaagtcca 540
attttttcga cgtctttcac ttgggccaca cacacaacta aagtaactag aagcgcaggc 600
tctaggaggc caccgttctg ttcacagtga agagggtgcg ctcaccgttg gtcgtgtccg 660
ctggaagccc cgcgtcaggc cgggagcggg acagagactc ttgctcaggg ccgttatccg 720
aactgatccg cttcccaccg cacccccaga gaaacccacc caaccccccta aacctaagaa 780
acccagactg cgcaaacctg caggaacaga gccatttccc cctaatgtgt gcttcaaacc 840
caccgaagcc caactgtaag caagaccagc gtgcccgccc tgcacgatac tgcttctccc 900
cgcagcagcg gctgccgatc tgggcagcgg gtgggtattc ctggggctcc gtggacgttg 960
agccgccgcg cgaaaccggc gccggctgga cctgcaaatc gccgcccggc cggcagggga 1020
cgccgcggac gcgagggcga ggtcggtcgc ccaggagggg gcgcgcgagg ccgcaggggc 1080
gggggcgcc gcctcacttg ccctgcgccc ctcccccgcg cgccctcctg gcgcggcggc 1140
cggcgaggcc cctgtgggag agggggcggg gacgaaacgg ccccgaggct cggagcgccg 1200
cgcggcggcg gcgcgagccc gagggggcgg ggaggcgcgg gcgggtgtgc gcgcgccggg 1260
cgtgggtgtg ggtgggggta accggcgcgg gcgccgagat agcgccgggc agagggaccc 1320
ggctaccctg gacagcgcat cgcc                                       1344

<210> 4
<211> 846
<212> DNA
<213> human

<400> 4
atggccttca gtgatcttac atcgaggact gtgcatcttt atgataattg gatcaaagat 60
gctgatccaa gagttgaaga ttggctcctc atgtcctcgc ctctgccaca aaccatcctc 120
ctaggattct atgtctattt tgtcacttcc ttgggaccaa agctcatgga aaatcgcaag 180
ccctttgaac tcaagaaagc aatgataacg tacaattttt tcatagtact cttttctgtg 240
tatatgtgtt atgagtttgt gatgtctggc tggggtatag gttattcatt tcgatgtgac 300
attgttgact attcacggtc acccacagct ttgaggatgg cacgtacctg ctggctttat 360
tacttctcca aatttattga gctattagat acgatctttt ttgttctgcg caagaaaaat 420
agccaagtga ctttccttca tgtattccat cataccatca tgccgtggac ctggtggttt 480
ggagtcaaat ttgctgcagg tggtttggga acattccatg cccttctaaa tacagctgta 540
catgtagtca tgtattccta ctatggactt tctgcattgg ggccagccta ccagaagtat 600
ttgtggtgga aaaaatattt gacatcatta cagcttgtcc agtttgttat tgtcgccatc 660
cacataagcc agttcttttt catggaggat tgcaagtatc agtttccagt ctttgcgtgc 720
atcattatga gttacagttt catgtttctg ctgctctttc tccatttttg gtaccgtgct 780
tacaccaaag gtcagaggtt gcccaaaact gtgaaaaatg gaacttgcaa aaacaaagat 840
```



```
aattga                                                            846

<210> 5
<211> 281
<212> PRT
<213> human

<400> 5
Met Ala Phe Ser Asp Leu Thr Ser Arg Thr Val His Leu Tyr Asp Asn
1               5                   10                  15

Trp Ile Lys Asp Ala Asp Pro Arg Val Glu Asp Trp Leu Leu Met Ser
            20                  25                  30

Ser Pro Leu Pro Gln Thr Ile Leu Leu Gly Phe Tyr Val Tyr Phe Val
        35                  40                  45

Thr Ser Leu Gly Pro Lys Leu Met Glu Asn Arg Lys Pro Phe Glu Leu
    50                  55                  60

Lys Lys Ala Met Ile Thr Tyr Asn Phe Phe Ile Val Leu Phe Ser Val
65                  70                  75                  80

Tyr Met Cys Tyr Glu Phe Val Met Ser Gly Trp Gly Ile Gly Tyr Ser
                85                  90                  95

Phe Arg Cys Asp Ile Val Asp Tyr Ser Arg Ser Pro Thr Ala Leu Arg
            100                 105                 110

Met Ala Arg Thr Cys Trp Leu Tyr Tyr Phe Ser Lys Phe Ile Glu Leu
        115                 120                 125

Leu Asp Thr Ile Phe Phe Val Leu Arg Lys Lys Asn Ser Gln Val Thr
    130                 135                 140

Phe Leu His Val Phe His His Thr Ile Met Pro Trp Thr Trp Trp Phe
145                 150                 155                 160

Gly Val Lys Phe Ala Ala Gly Gly Leu Gly Thr Phe His Ala Leu Leu
                165                 170                 175

Asn Thr Ala Val His Val Val Met Tyr Ser Tyr Tyr Gly Leu Ser Ala
            180                 185                 190

Leu Gly Pro Ala Tyr Gln Lys Tyr Leu Trp Trp Lys Lys Tyr Leu Thr
        195                 200                 205

Ser Leu Gln Leu Val Gln Phe Val Ile Val Ala Ile His Ile Ser Gln
    210                 215                 220

Phe Phe Phe Met Glu Asp Cys Lys Tyr Gln Phe Pro Val Phe Ala Cys
225                 230                 235                 240

Ile Ile Met Ser Tyr Ser Phe Met Phe Leu Leu Leu Phe Leu His Phe
                245                 250                 255

Trp Tyr Arg Ala Tyr Thr Lys Gly Gln Arg Leu Pro Lys Thr Val Lys
            260                 265                 270
```

Asn Gly Thr Cys Lys Asn Lys Asp Asn
275 280

<210> 6
<211> 2456
<212> DNA
<213> human

<400> 6
tgcgcctggc tgaacactac attttttta cttctttatt catgtattgt ctgtcatctc 60
caactagaat gaacgtatag tccctgagaa cggggaattt gttatctatt gaaacttcag 120
ggcctggaac atagcagcac tccagtattt gttaaataaa tgaatccatt tgagcttctg 180
catatttgaa atttcataag tatatataaa tggtaaattg tgatagactc aaaggctagt 240
atcattaggc aattgtctcc cgttcccaaa agacttccta agtctactaa atgatctgtt 300
tttaatatga aagcaaagtt atctaaaaga aaggagaaat ctttagtttt tttgacttcg 360
agattctttg caatttaagc tttttttttt tttttttttt tttttttttg cttttctttc 420
aatggacact ttcgaagttt tacataaaaa cattaaaacc tcttgtttaa tgtagtggga 480
ttaagctgcc gaaggcaatc cctacatgtg aggaaaatat gcttccgaca ccccaatttt 540
tttttctcc ctacccatcc tctctggtgg tcctgacgct cccagcccct ttttgtgttt 600
cttgattcca tgctgagaac tcgcaataca aactcaaagc ccacatttgt gaggtggttg 660
ggtcaggact gcaactaaaa atgattattg tttttaggt ttctggacag ttcaacacca 720
gcctttggtt ttgcctcaga agcagggaac ttctctaggc ccctattttg cctttcagct 780
attgatgatc caaatcatac cagcgattag gaggatcatt accagacaca aggccaggta 840
cgtttaaaaa ataaataaac caagcgcagg tgcacactcc gaacgctcat ccccaccccc 900
actttccaat ccaacagtag gtaacgagaa atgaattttc tagacttttt ttcctgcagc 960
agttgctgtt accagaaaca aagttagatg atatacaatc taatcttcat tgctctaaaa 1020
gtcctctccc catgcccccc aggctgcctc aattctctag tttcttattc cttataagca 1080
ggggatggag ctgaaccaag tcggccttcc cctcccaggg ccttctcctc ttggtctggc 1140
ttccatttca gatgcgaatt aaccctccca ataccctttc agaagcaagg agtccccttt 1200
ttctccgcct ccagcctcag ctaggttttc ctcatttcgg atttttctac agctcattcc 1260
caaatgagtc acgcatgacg acaatttcca ctctgctatg tcagcctgga gatgtccccc 1320
aagtgatggc atctgctctc ggaaagaaag gtcatcggtg ccacgaccag ccccgctaac 1380
ccagagcggc cggtgggccc cagtcccgag agtcagggcg cgcggcggag gcgaggccgg 1440
ggcggcctcc gccctcccgg ccgctccccc tcgcgccgcc ccggctcctc cctccggccc 1500
tcggcgggca cctggcggcg gcgggcaggg ggcggcgctg cgcgcgtcac gcggctgggt 1560
gggatagcgg gcaggtgaca cccggcggcc tcctcccctt tccaacccag tcggcggccg 1620
ggacagcagg ggccgctgtg aggagctccg cgctcgcgct gccagtcgcc gccctctctc 1680
ccgcgcgcgc ccggcgcttc ggctccgctc cctgtgcggt gagtgcgggg ttccaggccg 1740
gcgggcaggg gccaaacttt cccggcgcgc ggaggagaag agactgggga gggaggcaga 1800
gccgagggga acggcgtcgg gagtggccgg atggaggaac ttgggcgcgg cgcgcgagaa 1860
gtgggacccg ggtgcggggg ccccgggagc ggggccaggc cctccctggg ctcgggaggc 1920
gcttgggaag ttctgtcccc gctgcctgcg cgtggggagg accgaggccc ttttcgccgg 1980
agcgcggggc cgcggcgctc acctgcgcct tctcgggagc ccccacccgg cagcatcccg 2040
aagggaaggt cgggcccggt gggcgcgctg cggagcggag cctggactgg ggtcccgcgc 2100
ggcgctggcc ctgcggagcg gagcgggagg ggcagaggtg ctcgccggcg ggactgggag 2160
ggagaaggac ctgctcgacc ttggacgcgg aggtcatttt cccagctccg gggtctggcc 2220
tcgctagcca cccccccaaa ttccggagcc cctttctttc tgtttccttc cttccccttt 2280
gggcgctttt tttgctcccg cggccagatg aacttggggc gctgtccctt cggctccccg 2340
agccgcatcc tgtcttggtg gctgctgctg gccgggagga ggctgatgaa tacagagccg 2400
tggaacaggt cgtgccggag atggaaacag gaaagcctgt tgttttgtcg tcccag 2456

<210> 7
<211> 1411
<212> DNA
<213> human

<400> 7
gtgagccacc accgcggccg gtcccttcct cctttaaaaa tttttctccc agttcccact 60

ttttgtgggt tagaggcatc taaattgaat gaaagtaccc tttttggact actggggagg 120
tgggggatg ttctcagaag ggaattttc tttctggtcc taatatccac ctaattttta 180
aaagcagggc tccttattat tttgtaaagt ttacaattac atcattagat acttccatgt 240
ctcatatttc atttttccaa actcttgggg gaaatgagtg gagggatgga tggaatagaa 300
aatagttttt cctcttggag gctgagggcc cagtaggggt caacagtaca ttcagccctc 360
tcctcacata ttctgttcta cctacaagta cagcaagtaa agccaaattt ctcatgcatg 420
caaataaagt ttttgcattt ggccagtcgg tccagttctc ctgtcagctt ccttccccac 480
tctgcctctg ttcattaatc cccccttcc cggtacctaa accctccacc taacccagcc 540
ctttcttcca cttccggcta ctagcctctc tcgcctatcc actatcctca cactcagcat 600
cccctgtctg tacgagatta aggagctctg ccgtccgcag ggcctgggtt agcgtgaatc 660
taagccagag ctcccgggtg ggggtggggg taggggtggg ggtggtccca gaggtagggc 720
gaggaggtgg gaagcgtatt cccttcactg gtgatctcaa cgtagatttg cccggagttc 780
tcttgcaaga gagctggcag gttttactat ttcccaatcg tttactcgcc aagctctcgg 840
gtccacgcgc cgcggggatg cgccctgcaa gctgaaactt cattcaaagc aaggcggccc 900
acgaggttgg gcttagggga tctggatgac ctccaggcca cttcctttct ctctgcgccc 960
ttcccccact cttccaacca ccttcgctgt aaacaaaact gtcccccccg ggcggagaga 1020
ggtcgcgctc tttcgcacac tccctcgcca agggttaatt tctcaaatcg cacgaggggg 1080
aggagatttc cctgtagacg agtaaaaagg gtgatggaca aacgtgcggg cactaagacc 1140
gcaaggcatt catttcctcc tacggtggat gcggacgccg ggaggaggag agccccagag 1200
agaggagctg ggagcggagg cgcaggcaat gctcagccct ggatgtagct gagaggctgg 1260
gagaagagac gaccgctgga gaccgagcgg cgtggggaag acctaggggg gtgggtgggg 1320
gaagcagaca ggagaacact cgaaatcaag cgctttacag attattttat tttgtataga 1380
gaacacgtag cgactccgaa gatcagcccc a 1411

<210> 8
<211> 813
<212> DNA
<213> human

<400> 8
atggtcacag ccatgaatgt ctcacatgaa gtaaatcagc tgttccagcc ctataacttc 60
gagctgtcca aggacatgag gcccttttc gaggagtatt gggcaacctc attccccata 120
gccctgatct acctggttct catcgctgtg ggcagaact acatgaagga acgcaagggc 180
ttcaacctgc aagggcctct catcctctgg tccttctgcc ttgcaatctt cagtatcctg 240
ggggcagtga ggatgtgggg cattatgggg actgtgctac ttaccggggg cctaaagcaa 300
accgtgtgct tcatcaactt catcgataat tccacagtca aattctggtc ctgggtcttt 360
cttctcagca aggtcataga actcggagac acagccttca tcatcctgcg taagcggcca 420
ctcatcttta ttcactggta ccaccacagc acagtgctcg tgtacacaag ctttggatac 480
aagaacaaag tgcctgcagg aggctggttc gtcaccatga actttggtgt tcatgccatc 540
atgtacacct actacactct gaaggctgcc aacgtgaagc cccccaagat gctgcccatg 600
ctcatcacca gcctgcagat cttgcagatg tttgtaggag ccatcgtcag catcctcacg 660
tacatctgga ggcaggatca gggatgccac accacgatgg aacacttatt ctggtccttc 720
atcttgtata tgacctattt catcctcttt gcccacttct tctgccagac ctacatcagg 780
cccaaggtca aagccaagac caagagccag tga 813

<210> 9
<211> 270
<212> PRT
<213> human

<400> 9
Met Val Thr Ala Met Asn Val Ser His Glu Val Asn Gln Leu Phe Gln
1 5 10 15

Pro Tyr Asn Phe Glu Leu Ser Lys Asp Met Arg Pro Phe Phe Glu Glu
20 25 30

Tyr Trp Ala Thr Ser Phe Pro Ile Ala Leu Ile Tyr Leu Val Leu Ile

```
         35                  40                  45

Ala Val Gly Gln Asn Tyr Met Lys Glu Arg Lys Gly Phe Asn Leu Gln
    50                  55                  60

Gly Pro Leu Ile Leu Trp Ser Phe Cys Leu Ala Ile Phe Ser Ile Leu
65                  70                  75                  80

Gly Ala Val Arg Met Trp Gly Ile Met Gly Thr Val Leu Leu Thr Gly
                85                  90                  95

Gly Leu Lys Gln Thr Val Cys Phe Ile Asn Phe Ile Asp Asn Ser Thr
            100                 105                 110

Val Lys Phe Trp Ser Trp Val Phe Leu Leu Ser Lys Val Ile Glu Leu
        115                 120                 125

Gly Asp Thr Ala Phe Ile Ile Leu Arg Lys Arg Pro Leu Ile Phe Ile
    130                 135                 140

His Trp Tyr His His Ser Thr Val Leu Val Tyr Thr Ser Phe Gly Tyr
145                 150                 155                 160

Lys Asn Lys Val Pro Ala Gly Gly Trp Phe Val Thr Met Asn Phe Gly
                165                 170                 175

Val His Ala Ile Met Tyr Thr Tyr Tyr Thr Leu Lys Ala Ala Asn Val
            180                 185                 190

Lys Pro Pro Lys Met Leu Pro Met Leu Ile Thr Ser Leu Gln Ile Leu
        195                 200                 205

Gln Met Phe Val Gly Ala Ile Val Ser Ile Leu Thr Tyr Ile Trp Arg
    210                 215                 220

Gln Asp Gln Gly Cys His Thr Thr Met Glu His Leu Phe Trp Ser Phe
225                 230                 235                 240

Ile Leu Tyr Met Thr Tyr Phe Ile Leu Phe Ala His Phe Phe Cys Gln
                245                 250                 255

Thr Tyr Ile Arg Pro Lys Val Lys Ala Lys Thr Lys Ser Gln
            260                 265                 270


<210> 10
<211> 1937
<212> DNA
<213> human

<400> 10
gattagctgt caggctatat atggagccat caggaaccac tgaaggtttt tttttttttt 60
tttttttttg agacggagtc tcactctgtc acccaggctg gagtgcagtg gcacgatctc 120
tgctcactgc aagctctgcc tcccaggttc acgccattct cctgcctcag cctcccgagt 180
agctgggact acaggcgcct gccaccacgc ccggctaatt ttttgtattt tttagtagag 240
acggggtttg acggtgttag ccaggatggt ctcgatctcc tgacctcatg atctgcccgc 300
ctcggcctcc caaggtgctg ggattacagg cgtgaaccac cgtgcccggc cgaaccactg 360
aaggttttta agcaggaaag cagagctgtt ttctggatga gcaaacagaa agtagtggtt 420
ttccaagtac agtctgagac aacctatagg accagaatct ctgcagttga ggctcaggaa 480
tctggtaatc agccaggtat aggaactctt ttctgattgc aatgcagtga agagcagaag 540
```



```
cactgtatta gagaaagagg cagtgcaacc aggtaacgtg accaggtgag aagtgatgag 600
gtacagagac aaagagatgc acttttgagt cacttagatg gcactgatag gacttccact 660
acaccctcgc atagacagtg gctgaggttc aggaaataga gctggggttc ctacttggat 720
cctctggctc tagagcttta ctgcacatag ccatttatac ccacatcttg attttaatta 780
ttttatatct atgtttctta gcacttttg caaatttcca ccttatctca aactgccctc 840
aagccttgta tttctccttc gctttcataa aacctaggaa agaaataagg gacagccaag 900
taaaactttt aaaagtttta gaacatttat ttctttgggg ctggttacac aggcgagaaa 960
gaagtagatt tggttaggga gagaaaacaa caggccttgg ggagatacac tggctctccc 1020
cctccctaaa ccctaagagg cctccaggaa acctgaagac aataattcca gaagcccaga 1080
gggtgacccc atttcctctc tccatggtta ttactgtcag tctggagcag ttcaggaatt 1140
caggaaacta taaagaaacc acaacagcct caacaaccca aacatcaaca tcaacaacct 1200
caacaataaa actccttaaa attcatctcc ttccacccac tcacaaccgc agactcgaag 1260
ctaggaggtg gaagggacta cagaagctct gcgttgccca ggttagtatt tgctcatcac 1320
aggcctgggt ttcccaggat ctcagggagc ctggaaactg acgcctccat ttctgggtgg 1380
gagcaccaaa gcctaaggac acctttcctc tctcttcact gctaagcagg tcaagattaa 1440
agcaaaccga ggcaaaggcc acggttgaca gttccaaggg aacccgcaag gccgcacagg 1500
atggggtgga cgttttacgg gagaaaaggc tggggaagtg ggcgggcgat ggcctacgac 1560
gggacttggg gcggggtgtg cgaaacgcct ggcaggcggg cccttgagta tgaccaatca 1620
gaatgcggac tgcgtcccag gggcggagca gaggcgtatc ttggtcgaga ttggatagcg 1680
gcggggcgca ggaaagaggt cgcgccagcc cgggcaggca gctttgcaag tccgcgttat 1740
atatcgcagt ggctgcgccc gggatagctg gctgcgccgc cgcgcacatg cctaggttcg 1800
acgccctcct ccctttgccc aggagttcct tctgtcccgg ctctgttccg tctcgccccg 1860
aggttcacgc catcctcgga gccccagcct ttcacccagc gcctccaagc tttggacctt 1920
gacttctgca aaactag                                                 1937


<210> 11
<211> 945
<212> DNA
<213> human

<400> 11
atggggctcc tggactcgga gccgggtagt gtcctaaacg tagtgtccac ggcactcaac 60
gacacggtag agttctaccg ctggacctgg tccatcgcag ataagcgtgt ggaaaattgg 120
cctctgatgc agtctccttg gcctacacta agtataagca ctctttatct cctgtttgtg 180
tggctgggtc caaaatggat gaaggaccga gaaccttttc agatgcgtct agtgctcatt 240
atctataatt ttgggatggt tttgcttaac ctctttatct tcagagagtt attcatggga 300
tcatataatg cgggatatag ctatatttgc cagagtgtgg attattctaa taatgttcat 360
gaagtcagga tagctgctgc tctgtggtgg tactttgtat ctaaaggagt tgagtatttg 420
gacacagtgt tttttattct gagaaagaaa aacaaccaag tttctttcct tcatgtgtat 480
catcactgta cgatgtttac cttgtggtgg attggaatta agtgggttgc aggaggacaa 540
gcattttttg gagcccagtt gaattccttt atccatgtga ttatgtactc atactatggg 600
ttaactgcat ttggcccatg gattcagaaa tatctttggt ggaaacgata cctgactatg 660
ttgcaactga ttcaattcca tgtgaccatt gggcacacgg cactgtctct ttacactgac 720
tgccccttcc ccaaatggat gcactgggct ctaattgcct atgcaatcag cttcatattt 780
ctctttctta acttctacat tcggacatac aaagagccta agaaaccaaa agctggaaaa 840
acagccatga atggtatttc agcaaatggt gtgagcaaat cagaaaaaca actcatgata 900
gaaaatggaa aaaagcagaa aaatggaaaa gcaaaaggag attaa                  945


<210> 12
<211> 314
<212> PRT
<213> human

<400> 12
Met Gly Leu Leu Asp Ser Glu Pro Gly Ser Val Leu Asn Val Val Ser
1               5                   10                  15

Thr Ala Leu Asn Asp Thr Val Glu Phe Tyr Arg Trp Thr Trp Ser Ile
```

```
                20                  25                  30

Ala Asp Lys Arg Val Glu Asn Trp Pro Leu Met Gln Ser Pro Trp Pro
        35                  40                  45

Thr Leu Ser Ile Ser Thr Leu Tyr Leu Leu Phe Val Trp Leu Gly Pro
    50                  55                  60

Lys Trp Met Lys Asp Arg Glu Pro Phe Gln Met Arg Leu Val Leu Ile
65                  70                  75                  80

Ile Tyr Asn Phe Gly Met Val Leu Leu Asn Leu Phe Ile Phe Arg Glu
                85                  90                  95

Leu Phe Met Gly Ser Tyr Asn Ala Gly Tyr Ser Tyr Ile Cys Gln Ser
            100                 105                 110

Val Asp Tyr Ser Asn Asn Val His Glu Val Arg Ile Ala Ala Ala Leu
        115                 120                 125

Trp Trp Tyr Phe Val Ser Lys Gly Val Glu Tyr Leu Asp Thr Val Phe
    130                 135                 140

Phe Ile Leu Arg Lys Lys Asn Asn Gln Val Ser Phe Leu His Val Tyr
145                 150                 155                 160

His His Cys Thr Met Phe Thr Leu Trp Trp Ile Gly Ile Lys Trp Val
                165                 170                 175

Ala Gly Gly Gln Ala Phe Phe Gly Ala Gln Leu Asn Ser Phe Ile His
            180                 185                 190

Val Ile Met Tyr Ser Tyr Tyr Gly Leu Thr Ala Phe Gly Pro Trp Ile
        195                 200                 205

Gln Lys Tyr Leu Trp Trp Lys Arg Tyr Leu Thr Met Leu Gln Leu Ile
    210                 215                 220

Gln Phe His Val Thr Ile Gly His Thr Ala Leu Ser Leu Tyr Thr Asp
225                 230                 235                 240

Cys Pro Phe Pro Lys Trp Met His Trp Ala Leu Ile Ala Tyr Ala Ile
                245                 250                 255

Ser Phe Ile Phe Leu Phe Leu Asn Phe Tyr Ile Arg Thr Tyr Lys Glu
            260                 265                 270

Pro Lys Lys Pro Lys Ala Gly Lys Thr Ala Met Asn Gly Ile Ser Ala
        275                 280                 285

Asn Gly Val Ser Lys Ser Glu Lys Gln Leu Met Ile Glu Asn Gly Lys
    290                 295                 300

Lys Gln Lys Asn Gly Lys Ala Lys Gly Asp
305                 310


<210> 13
<211> 2000
<212> DNA
```



```
<213> human

<400> 13
ggaaatacct gaagctgttt taacaatttc tccttgtatt aagtattatg ctgcagtttt 60
gcgtgtgtga atggaagtat gggtagagat ctgttctccc taaaaactcc aggattccac 120
aatatagaaa tagtaatcaa atttttaggt gaagctcgaa ctaatccgaa ctttgttaga 180
tcatcactgt aaatgaatgg gtatttatcc actccctaaa tgaagagact tgactggatt 240
tctttttttt atatagctac tagaatctgt tacacataat ttaggattga gacttgagaa 300
attgtcattc caatccagaa aactttagat ttgcaaatat atttgacaaa ttaataaatt 360
aacattttat ttggttaatt tcaagaatag ggcatttaaa gaagtctgtg tttgctttag 420
ttcggcaata aagttcctgc cactcacaat aatccttatt attctctgaa agacatgtta 480
tatttttgtc atcataaata tttattaatt actgtttata gcactgggtt aggtactcat 540
caagcaacca aaaataattc ttaccatcta ggatgcttcc aatataaaat atagacaata 600
tataaccagg tcaattggga aatagatcat ttcagtatga taaaagatag tattcacatt 660
aacagtgtga aagggcagga acaataagac acttgactca ctggtcttta aaatgtagca 720
tccaaaatga gcaagtggag aaaaggttaa acaagtaggt gacacattta aaaaacaagt 780
agatgaaagg actattctca aaaatcttgt tttatgtgag aaaccatcaa attatgaatt 840
ccaagtactg tatttttttt acttttcaag ggtaggctct cctatacctt atctaaacaa 900
tttttcaaaa tagccacaat tactttgttt tcctctctac actaaattgc cctttgcctc 960
ttgagcgatt atctttttca gattcacctc aacttcttca ggttcaagcg gacttcacct 1020
gtaagcccct ctcggttctc cctcttctct gaactactaa tggcctaatt tagcacaatt 1080
atattgcttt gttcattcca tgtatagtaa aagagtctac aaaacacatg caagcattca 1140
tgcaattata tgttgatttg ttcatgggtc gaccccaaag tctattctcc atcgctgaag 1200
catggaagac aaataccctt cacttcttca gaggcataac acatgcactt ctcttgtcat 1260
ggtgacaggc atgtgctggt ggaggtcaaa gaaacaggaa cacaagtgaa atcgaggtga 1320
gtgtcaggta aggaccaaag caccacgcct acctcatctt tgcccacaga acacccattc 1380
ttcccgtgtc ctgtttccca ggacgtatcc ggggcggata agaaatcacc cgtggggagg 1440
cggtgaactc ctccgcaggg gccgatgccc gggacagggg cggggaaggc taatgaggcg 1500
acttgtgcgg ggaggggcca aggaggagcc caggtgtccc gctcccgctc gacggcgcgc 1560
gcctgcgcga gcccagttgg cgtcgcaccc ttgagcgcag catccctacg ccagcgagtc 1620
ccaatactag ggagggaggg agggaggagg ggcggccggc cccccgcccc cgcgcgcggc 1680
cacgtgacgc cggctgagga gattggaggg gcggctgcgc gaggctgcag actggtgcag 1740
cgcactgtgc tggcggctgg gcctcctcca cctcctcgtc tttctcccgg gaaccttgac 1800
gacgccttcc gcttggccct gccttctgcc gcatccccgc cgccgcggcg ccttgaggag 1860
caggagaaga cgcagccggg ccgccgccgt tagaggggtt cccggccgcc gctcgccccg 1920
tcggccgcca ccgcctccgg ggtcagccct ctctctgggt ctccgctttc tcctgccgcc 1980
agcgcccgct catcgccgcg                                              2000


<210> 14
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 14
cacgcgggta ccaggatgga ggctgttgtg aac                               33


<210> 15
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 15
```

atatcacgat gcggccgctc agttggcctt gaccttggc 39

<210> 16
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 16
atatcacgat gcggccgcca gttggccttg accttggc 38

<210> 17
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 17
gtaacaggag tatgggaagg ca 22

<210> 18
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 18
ttggactcac actgctgtct ct 22

<210> 19
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 19
gtgtggcacc aaaataagag tg 22

<210> 20
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 20

cacgcgggat cccaaatgga acattttgat gcatcac 37

<210> 21
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 21
atatcacgat gcggccgctc aatccttccg cagcttcc 38

<210> 22
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 22
atatcacgat gcggccgcca atccttccgc agcttcc 37

<210> 23
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 23
cacgcgggat ccatcatgga acatctaaag gcc 33

<210> 24
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 24
atatcacgat gcggccgctt attgtgcttt cttgttcatc actcc 45

<210> 25
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 25

atatcacgat gcggccgctt ttgtgctttc ttgttcatca ctcc 44

<210> 26
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 26
gccagcctac cagaagtatt tg 22

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 27
gcgcaagaaa aatagccaag 20

<210> 28
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 28
aatgatgcac gcaaagactg 20

<210> 29
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 29
cacgcgggat ccctgatgaa tacagagccg tgg 33

<210> 30
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 30

atatcacgat gcggccgctc aattatcttt gtttttgcaa gttcc 45

<210> 31
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 31
cacgcgggat ccctgatgga aaagcccatt aatattc 37

<210> 32
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 32
atatcacgat gcggccgctc aattatcttt gtttttgcaa gttcc 45

<210> 33
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 33
atatcacgat gcggccgcca attatctttg tttttgcaag ttcc 44

<210> 34
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 34
cacgcgggat ccaaaaatga acatgtcagt gttgacttta caag 44

<210> 35
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 35

atatcacgat gcggccgcct attcagcttt cgttgttttc ctc 43

<210> 36
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 36
atatcacgat gcggccgcca ttcagctttc gttgttttcc tc 42

<210> 37
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 37
cacgcgggat ccaaaaatgg tcacagccat gaatgtctc 39

<210> 38
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 38
atatcacgat gcggccgctc actggctctt ggtcttggc 39

<210> 39
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 39
atatcacgat gcggccgcca ctggctcttg gtcttggc 38

<210> 40
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 40

cacgcgggat ccaaaaatgg ggctcctgga ctcggagc 38

<210> 41
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 41
atatcacgat gcggccgctt aatctccttt tgcttttcca tttttctgc 49

<210> 42
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 42
atatcacgat gcggccgctt atctcctttt gcttttccat ttttctgc 48

<210> 43
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 43
ggaagatctt acaggctcgt gaggcttccc tcccg 35

<210> 44
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 44
ggaagatctc cggcaggagg gaccaaggct 30

<210> 45
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 45

ggaagatctc gagggtgggc ttctgccacc c 31

<210> 46
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 46
ggaagatctc ttttagccca aggggcggca gc 32

<210> 47
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 47
ggaagatctt tcgtgtgaat ttccttcaag tctc 34

<210> 48
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 48
ggaagatctt gatccgcagc ggctgtg 27

<210> 49
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 49
ggaagatctc cggtacctac agttactcac tctgc 35

<210> 50
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 50

ggaagatctg gcgatgcgct gtccagggta 30

<210> 51
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 51
cgacgcgttg cgcctggctg aacactac 28

<210> 52
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 52
ggaagatctc tgggacaaac aacaggc 27

<210> 53
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 53
ccgctcgagg tgagccacca ccgcggcc 28

<210> 54
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 54
ccgctcgagt ggggctgatc ttcggagtcg c 31

<210> 55
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 55

```
ccgagctcga ttagctgtca ggctatatat ggagcc                   36


<210> 56
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 56
ccgagctcct agtttgcaga aggtccaaag c                        31


<210> 57
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 57
ccgagctcgg aaatacctga agctgtttta ac                       32


<210> 58
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 58
ccgagctccg cggcgatgag cgggc                               25
```